



ARL-TR-7729 • JULY 2016



# **US Army Research Laboratory (ARL) Robotics Collaborative Technology Alliance 2014 Capstone Experiment**

**by Marshal Childers, Craig Lennon, Barry Bodt, Jason Pusey,  
Susan Hill, Rick Camden, Jean Oh, Robert Dean, Terence  
Keegan, Chip Diberardino, Sisir Karumanchi, Bertrand  
Douillard, Nikhil Gupta, Camilo Ordonez, Jacob Shill, Emmanuel  
Collins, Jonathan Clark, Daniel Barber, Jeff Duperret, Garrett  
Wenger, Menglong Zhu, Sangdon Park, Daniel Koditschek,  
Aaron Johnson, Arne Suppe, and Luis Navarro-Serment**

Approved for public release; distribution is unlimited.

## **NOTICES**

### **Disclaimers**

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

Citation of manufacturer's or trade names does not constitute an official endorsement or approval of the use thereof.

Destroy this report when it is no longer needed. Do not return it to the originator.



# **US Army Research Laboratory (ARL) Robotics Collaborative Technology Alliance 2014 Capstone Experiment**

**by Marshal Childers, Craig Lennon, Barry Bodt, Jason Pusey, Susan Hill, and  
Rick Camden**  
*Vehicle Technology Directorate, ARL*

**Jean Oh**  
*National Robotics Engineering Center, Pittsburgh, PA*

**Robert Dean, Terence Keegan, and Chip Diberardino**  
*General Dynamics Land Systems, Westminister, MD*

**Sisir Karumanchi and Bertrand Douillard**  
*Jet Propulsion Laboratory, Pasadena, CA*

**Nikhil Gupta, Camilo Ordonez, Jacob Shill, Emmanuel Collins, and Jonathan Clark**  
*Florida State University, Tallahassee, FL*

**Daniel Barber**  
*University of Central Florida, Orlando, FL*

**Jeff Duperret, Garrett Wenger, Menglong Zhu, Sangdon Park, and Daniel Koditschek**  
*University of Pennsylvania, Philadelphia, PA*

**Aaron Johnson, Arne Suppe, and Luis Navarro-Serment**  
*Carnegie Mellon University, Pittsburgh, PA*

Approved for public release; distribution is unlimited.

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188		
<p>Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p> <p><b>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</b></p>					
1. REPORT DATE (DD-MM-YYYY) July 2016		2. REPORT TYPE Final		3. DATES COVERED (From - To) 1 December 2013–31 December 2014	
4. TITLE AND SUBTITLE US Army Research Laboratory (ARL) Robotics Collaborative Technology Alliance 2014 Capstone Experiment			5a. CONTRACT NUMBER		
			5b. GRANT NUMBER		
			5c. PROGRAM ELEMENT NUMBER		
6. AUTHOR(S) Marshal Childers, Craig Lennon, Barry Bodt, Jason Pusey, Susan Hill, Rick Camden, Jean Oh, Robert Dean, Terence Keegan, Chip Diberardino, Sisir Karumanchi, Bertrand Douillard, Nikhil Gupta, Camilo Ordonez, Jacob Shill, Emmanuel Collins, Jonathan Clark, Daniel Barber, Jeff Duperret, Garrett Wenger, Menglong Zhu, Sangdon Park, Daniel Koditschek, Aaron Johnson, Arne Suppe, and Luis Navarro-Serment			5d. PROJECT NUMBER		
			5e. TASK NUMBER		
			5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) US Army Research Laboratory ATTN: RDRL-VTA Aberdeen Proving Ground, MD 21005-5066			8. PERFORMING ORGANIZATION REPORT NUMBER  ARL-TR-7729		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)			10. SPONSOR/MONITOR'S ACRONYM(S)		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)		
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.					
13. SUPPLEMENTARY NOTES					
14. ABSTRACT  This report highlights the capabilities demonstrated during the US Army Research Laboratory Robotics Collaborative Technology Alliance Capstone Experiment that took place during October 2014. The report succinctly presents the activities of the event and provides references for further reading on the specifics of those activities. Four capabilities were evaluated as part of distinct Integrated Research Assessments (IRA): Human Robot Interaction Modalities, Semantic Navigation and Perception, Search and Observation of Doorways, and Search and Grasping of Objects in an Indoor Environment. Five task-based assessments (TBAs) using various platforms were also conducted during this timeframe, which consisted of the following capabilities: Bracing to Reach and Grasp an Object, Detection and Climbing of Stairs, Leaping over a Span, Dynamically Feasible Motion Planning, and Terrain Aware Motion Planning. The results from the IRA and TBA events characterized performance and highlighted areas that call for continued and focused efforts in order to enable advancements in intelligence-based capabilities sufficient for the teaming of autonomous systems with Soldiers.					
15. SUBJECT TERMS robotics, experimentation, technology assessment, perception, semantic navigation, manipulation, legged mobility					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 62	19a. NAME OF RESPONSIBLE PERSON Marshal Childers
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified			19b. TELEPHONE NUMBER (Include area code) 410-278-7996

## Contents

---

<b>List of Figures</b>	<b>v</b>
<b>List of Tables</b>	<b>vii</b>
<b>1. Introduction</b>	<b>1</b>
<b>2. Background</b>	<b>1</b>
2.1 Robotics CTA	1
2.2 Assessment During First Five Years of Robotics CTA	2
<b>3. Capstone Experiment</b>	<b>3</b>
3.1 Purpose	3
3.2 Method	3
<b>4. Capstone Experiment Integrated Research Assessments</b>	<b>5</b>
4.1 IRA: Semantic Navigation and Perception	6
4.1.1 Common World Model	6
4.1.2 Perception Method	7
4.1.3 Navigation Method	9
4.2 IRA: Search and Observe Doorways	11
4.2.1 Search	12
4.2.2 Observe	12
4.3 IRA: Human Robot Interaction Modalities	14
4.4 End-to-End Scenarios	15
4.4.1 Evaluating the System	16
4.4.2 Design for the Experiment	16
4.4.3 Evaluation Criteria	18
4.4.4 Conclusions	19
4.5 IRA: Indoor Search and Grasp	19
4.5.1 Mission and Site	19
4.5.2 Physical System	21
4.5.3 Experimental Design	22
4.5.4 Metrics	23
4.5.5 Results Commentary	23

<b>5. Capstone Experiment Task-Based Assessments (TBA)</b>	<b>24</b>
5.1 TBA: Self-Anchored Reaching (JPL)	24
5.1.1 Introduction: Overview	24
5.1.2 Experimental Design and Task-Based Assessment Criteria and Metrics	26
5.1.3 Results	27
5.1.4 Conclusion	30
5.2 TBA: Stair-Climbing with XRHex	30
5.2.1 Introduction: Autonomous Stair Climbing Overview	30
5.2.2 Experimental Protocol	30
5.2.3 Results and Analysis	32
5.2.4 Conclusion	34
5.3 TBA: Gap-Crossing with Canid Quadruped	34
5.3.1 Introduction	34
5.3.2 Experimental Protocol	34
5.3.3 Results and Analysis	36
5.3.4 Conclusion	38
5.4 TBA: Efficient Motion with Dynamic and Power Models	38
5.4.1 Description of Capability	38
5.4.2 Description of Experiments	40
5.4.3 Results and Analysis	40
5.4.4 FSU-BOT	41
5.4.5 HUSKY ROBOT	42
5.4.6 Comments	42
5.5 TBA: Improved Contact Sensors for Terrain Classification	43
5.5.1 Description of Capability	43
5.5.2 Platform configuration	43
5.6 Results	44
<b>6. Conclusion</b>	<b>45</b>
<b>7. References</b>	<b>47</b>
<b>List of Symbols, Abbreviations, and Acronyms</b>	<b>51</b>
<b>Distribution List</b>	<b>52</b>

## List of Figures

Fig. 1	Combined arms collective training facility.....	4
Fig. 2	Husky UGV configuration for semantic perception and navigation IRA.....	6
Fig. 3	A front view of a building, with pixels colored according to semantic label. The building, car, and traffic barrel are labeled with text, and the green triangle shows the position and orientation of the robot. ....	7
Fig. 4	Semantic object detection, with actual objects on the left, and the labeled points of the 3-D point cloud on the right .....	8
Fig. 5	ADPM object detection finding gas pumps and a traffic barrel (outlined in pink).....	8
Fig. 6	Features in the world model produced by the semantically labeled image in Fig. 3. The blue walls are observed, and the grey walls are predicted. The robot is represented as a red arrow, and the planned path as a green curve leading to the predicted traffic barrel. ....	10
Fig. 7	The rear of the bar, from the robot's perspective, at the completion of run 15. Perceived walls are dark blue, and predicted walls are grey. Also shown is a fire hydrant and the misperceptions of a traffic barrel and car. ....	10
Fig. 8	Doors detected on the church.....	11
Fig. 9	Correctly placed detection boxes from the pedestrian detection algorithm .....	13
Fig. 10	These images show representations of pedestrians on the a) metric and b) symbolic level.....	13
Fig. 11	The tablet and glove used for and HRI interface. The left side of the tablet's screen displays a view combining an a priori map from OpenStreetMap (open) with objects in its own world model. The right side, from top to bottom, shows a view through the robot's camera, the command it is executing, and the activity it is trying to perform. ....	14
Fig. 12	Speech is used to command the robot to maneuver in the environment .....	14
Fig. 13	Test site positions referenced in Table 1. The position of the HMMWV in runs 7–9 and 14 is denoted as H. ....	17
Fig. 14	Number of runs achieving each score .....	19
Fig. 15	FTIG High Bay experimental area.....	20
Fig. 16	Schematic of High Bay experimental area.....	21
Fig. 17	Experimental system components.....	22
Fig. 18	The surrogate platform and the 360° field of view sensor head .....	25
Fig. 19	Extended reaching (~5 ft) in a bracing maneuver.....	25

Fig. 20	Multiple navigation runs with the a goal of {0,0} and different start locations .....	28
Fig. 21	Example of large displacement between 2 LADAR scans when running the perception pipeline at low rate. The reference scan is shown in red, the current scan that will be aligned to the reference scan is colored by segmentation. As can be seen, the main structures in the scene are off by about 45° between the 2 scans.....	29
Fig. 22	Experimental setup. An RGBD camera and Mac Mini are mounted on XRHex. Off screen, an operator is holding a joystick used to drive the robot when away from stairs.....	31
Fig. 23	Canid leaping a span .....	35
Fig. 24	Canid leap at drainage ditch.....	35
Fig. 25	Canid after failing to cross the drainage ditch (left) and the ground scuffmarks from insufficient traction in Canid's rear legs while jumping (right) .....	37
Fig. 26	FSU-BOT equipped with JPL visual odometry and FSU low-level data logging. In addition, the robot has a bay to modify the payload during experimentation using steel slabs.....	39
Fig. 27	Husky robot equipped with JPL's visual odometry and FSU's low-level data logging system (located in the lower bay of the robot). The computer runs the real-time operating system QNX and logs motor currents, IMU data, and odometry. ....	39
Fig. 28	Comparison of minimum distance and minimum energy trajectories. (E represents energy and D represents distance). The execution of the minimum distance trajectory resulted in an obstacle collision. ....	40
Fig. 29	Husky robot executing energy-efficient motion planning on asphalt..	41
Fig. 30	Energy prediction error for the different surfaces. Asphalt is represented in red, concrete in blue, and grass in green. ....	42
Fig. 31	(left) The experimental setup for terrain classification using PreSRS on the Hopper. (right) A computer-aided design schematic of the Hopper with PreSRS attached to the bottom of the robot foot. ....	44
Fig. 32	Plots of terrain classification accuracy vs. sensor damage attained from the damaged image sets (dashed line) and the repaired image sets (red line). The accuracy drops below 90% at 13% damage with no repairing, and at 90% damage with repair. ....	45



## List of Tables

---

Table 1	End-to-end scenario run details .....	16
Table 2	Scores, weather, and the reason given by the evaluator for scores less than 100.....	18
Table 3	Self-anchored reaching run details.....	26
Table 4	End-to-end navigation error, at the end of plan, of the robot and commanded goal locations.....	28
Table 5	xRhex Experimental setup mass measurements .....	31
Table 6	Stair measurements and robot measurements taken from depth camera .....	32
Table 7	xRhex stair-climbing experimental results .....	33
Table 8	Gap crossing data detailing Canid leaps from a Pelican case to the bank ledge .....	36
Table 9	Energy Prediction Errors (negative signs represent over estimation)..	42
Table 10	Terrain classification accuracies .....	44

INTENTIONALLY LEFT BLANK.

## **1. Introduction**

---

The purpose of this report is to highlight the capabilities demonstrated during the US Army Research Laboratory (ARL) Robotics Collaborative Technology Alliance (RCTA) Capstone Experiment that took place during October 2014. The document succinctly presents the activities of the event and provides references for further reading on the specifics of those activities. Given that the experiment consisted of numerous technologies, platforms, and researchers, the reports on specific experiments will be published in various conferences and articles and can stand on their own. This report is an opportunity to pull together all of these activities in one place so that the reader can appreciate the overarching program goals and understand the progress to date in realizing those goals through the preparation, integration, and conduct of relevant, structured experimentation.

## **2. Background**

---

### **2.1 Robotics CTA**

---

The RCTA is a fundamental research program that began in 2010 and enables Government, industrial, and academic institutions to address research and development required to enable the deployment of future military unmanned ground vehicle (UGV) systems ranging in size from man-portables to ground combat vehicles. Currently the consortium consists of the following partners: Carnegie Mellon University, General Dynamics Land Systems (Integration Lead), Florida State University (FSU), the Jet Propulsion Laboratory (JPL), Massachusetts Institute of Technology (MIT), QinetiQ North America, the University of Central Florida, and the University of Pennsylvania. The program is investing basic and applied research funding in 4 interdependent focus areas: Perception, Intelligence, Human-Robot Interaction, and Dexterous Manipulation and Unique Mobility. Long-term payoff for these efforts is envisioned by the following statement, which appears in the RCTA fiscal year 2012 Annual Program Plan (APP): “The future for unmanned systems lies in the development of highly capable systems, which have a set of intelligence based capabilities sufficient to enable the teaming of autonomous systems with Soldiers” (RCTA 2012). To realize these capabilities, progress must be made in the ability of the robot to think, look, talk, move, and work. The experiments described in this report include efforts that will enable advancements in all of these abilities.

The work presented here is made possible through experience attained from years of experimentation in relevant environments and scenarios with unmanned autonomous robots. A product of this experience is a technology assessment process that yields valuable information to the researchers, managers, and stakeholders of the program. Technology assessments are experiments designed by the government and planned and executed in cooperation with the members of the RCTA consortium. A detailed example of one ARL technology assessment is available in the journal article “Assessing Unmanned Ground Vehicle Tactical Behaviors Performance” (Childers et al. 2011).

Robotics by its nature involves the integration of technologies that enable a capability that exceeds the sum of the parts. During the course of developing the skills to assess these technologies, we learned that these integrated systems must be evaluated using a plan that accounts for that increased capability. The result is an integrated research assessment (IRA) wherein the performance of multiple integrated technologies is evaluated. A rudimentary example would be the coupling of perception with navigation to enable the robot to maneuver in an environment. This approach has an advantage in that the forced interaction of the collected research components often brings system-level considerations to light, which would not have otherwise been identified this early in the research and development process.

While integrating multiple technologies is often necessary to appreciate a capability, progress of a research program such as the RCTA varies in pace and maturity. In some instances the breadth of the program and goals for integration make it clear that some technologies will not find their way into an IRA in the near-term. In these instances we have found value in applying the assessment process to technologies on a task-based level. In a task-based assessment (TBA) there is usually some level of integration required to evaluate performance in a given environment or scenario, but the capability is limited and not readily integrated into an IRA. Assessment data at this level help the researcher to identify things that require attention and accelerates development.

## **2.2 Assessment During First Five Years of Robotics CTA**

---

During the first 5 years of the program, the RCTA conducted a number of IRAs and TBAs. In August 2011 we conducted a baseline assessment of autonomous UGV perception and intelligence. The 2-fold purpose of this event was to initiate the experimental component of the program and to evaluate current primitive robotic vehicle behaviors in a relevant environment (Bodt 2011, Bodt et al. 2012). It involved a Talon-based platform that could maneuver through an unknown environment to perform object detection and mapping. This capability represented

the state of the art for an autonomous unmanned vehicle to meet the look and move requirements. In November 2011 we conducted IRA1, which involved the detection and tracking of moving pedestrians from a stationary vehicle, which provided data for perception of moving objects. In May 2012 IRA2 assessed the ability of a Packbot-based autonomous vehicle to trench for buried wires. This event provided the opportunity to integrate a manipulator arm on the robot and obtain data on the ability of the system to perform work in a relevant environment (Bodt et al. 2013). In October 2012 we conducted IRA3 to evaluate the ability of a UGV to detect and classify objects using semantic perception techniques (Lennon et al. 2013). In early 2013 the state of the technologies would not enable the fourth planned IRA to be conducted; however, a number of TBAs were conducted in the areas of autonomous grasping, terrain dependent motion planning, and whole body dynamic manipulation (Murphy et al. 2013). In December 2013 IRA5 addressed the ability to perform semantic perception and navigation in a Mounted Operations in Urban Terrain environment (Lennon 2015a). During the IRA5 timeframe there were also TBAs performed on natural language translation and the performance of a gesture glove (Harris and Barber 2014).

### **3. Capstone Experiment**

---

#### **3.1 Purpose**

---

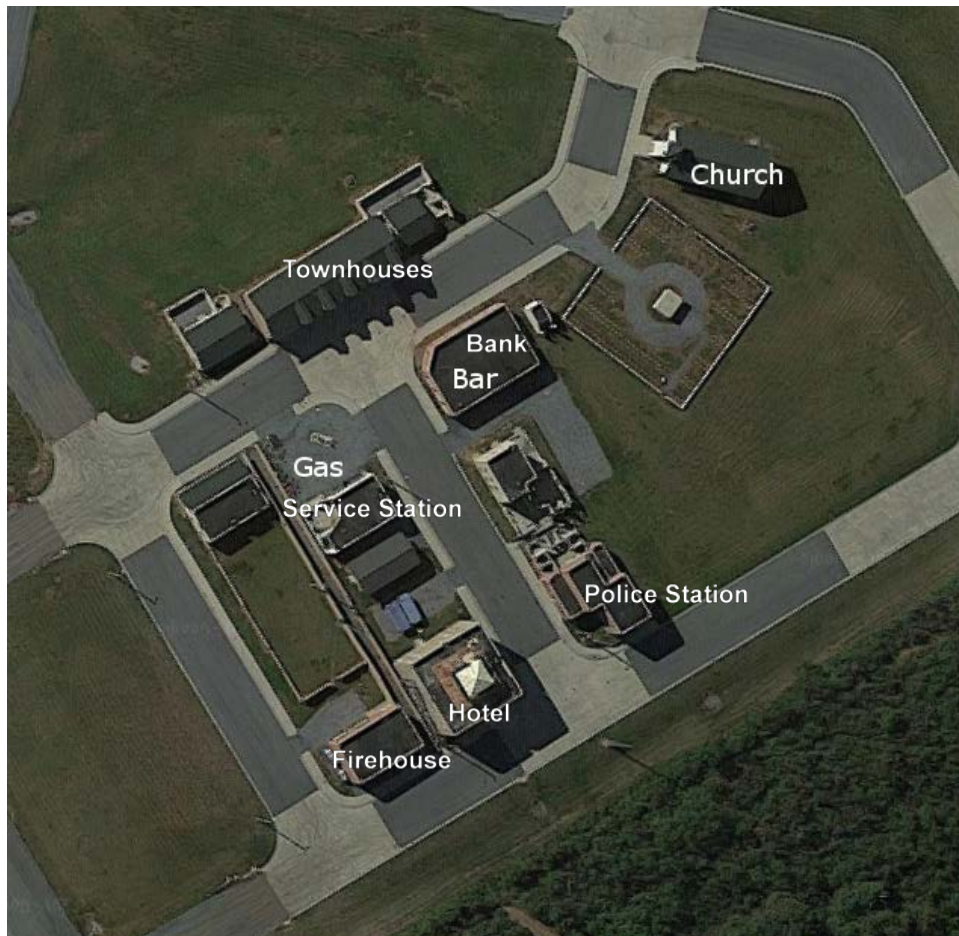
The RCTA capstone experiment took place in October 2014, approximately the mid-point of the program timeline, and represents progress achieved in the research thrust areas. The event was primarily held at Fort Indiantown Gap (FTIG), PA, at the Combined Arms Collective Training Facility (CACTF). Four capabilities were evaluated as part of distinct IRAs: Human Robot Interaction Modalities, Semantic Navigation and Perception, Search and Observation of Doorways, and Search and Grasping of Objects in an Indoor Environment. Data were also achieved for a fifth IRA, which consisted of stringing together the first 3 listed capabilities in a series of end-to-end runs. Five TBAs using various platforms were also conducted during this timeframe, which consisted of the following capabilities: Bracing to Reach and Grasp an Object, Detection and Climbing of Stairs, Leaping over a Span, Dynamically Feasible Motion Planning, and Terrain Aware Motion Planning.

#### **3.2 Method**

---

The CACTF consists of 9 full-scale buildings with paved streets and concrete curbs and sidewalks (Fig. 1). The experimental design leveraged the available features and terrain to evaluate the capabilities in a relevant manner. The church building

and the surrounding lawn was a focal point for many of the IRA data collections. The room to maneuver, ability to approach features of the church exterior from multiple angles, and the degree to which this building was apart from the other structures made it attractive for detecting and navigating among various objects (Semantic Navigation and Perception); finding a doorway, positioning the robot to observe the doorway, and detect pedestrians exiting the doorway (Search and Observe Doorways); and exercising multiple modes of commanding and interacting with a robot (Human Robot Interaction Modalities). Portions of the data collections for some IRAs were conducted in the vicinity of additional buildings and features to ensure that the data would exclude biases for particular areas of the CACTF and include features that the church building could not provide. For example, one of the features used in Semantic Navigation and Perception is a gas pump that is located in the vicinity of the service station and bar/bank buildings.



**Fig. 1** Combined arms collective training facility

Each IRA is the result of the process required to develop an experimental design and plan that will provide a reasonable ability to exercise the capabilities over a number of variables. The purpose of these efforts is not only to see how well something works when used in the manner and environment for which it was designed but to push the performance limits in a number of ways to reveal strengths and weaknesses of the current instantiation. Through collaboration with the researchers to understand the capabilities and underlying technologies, ARL was able to independently construct unique data collection protocols for each capability. Some designs benefited from sufficient available features and objects in the environment to provide a balanced data set. In numerous cases, to achieve a reasonable data set the designs required adjustment to accommodate the given abilities of the platforms and sensors in a particular environment.

#### **4. Capstone Experiment Integrated Research Assessments**

---

##### **Mission Description:**

The integrated assessments were based on a scenario in which the robot is told to screen the back door of a building. The scenario begins when the robot receives instructions, in a structured language, through a human-robot interaction (HRI) interface. These instructions give positions to which the robot should navigate and objects to be used as landmarks. If the robot successfully navigates to the correct position, it will detect and orient toward a door on the building, subsequently detecting and tracking pedestrians exiting through the door. If it is not successful, the HRI interface allows the operator to control the robot or to clarify ambiguous commands. For example, if the robot finds more than one door, or sees more than one possible goal building, it will give the operator an opportunity to assist it in choosing the correct one. This screening mission was executed as 17 complete runs intended to explore the combined capabilities of the system and as a larger number of runs testing parts of the mission in more structured, preliminary experiments. The mission is decomposed into a sequence of actions (e.g., navigate, search, observe), where each action has its own goal. This goal is generally the precondition of the next action in the mission plan. These experiments evaluated semantic navigation and perception, door detection, pedestrian detection and tracking, and human-robot interaction. In the following sections we summarize each IRA and also present performance evaluation of the first 3 listed capabilities as they appeared in the End-to-End Scenarios.

## 4.1 IRA: Semantic Navigation and Perception

---

### Robotic Platform:

The robot used in the integrated assessment is a Clearpath Husky, equipped with the General Dynamics XR 3-D (3-dimensional) laser detection and ranging (LADAR) sensor, Bumblebee stereo camera, and Adonis camera as shown in Fig. 2. The XR LADAR sensor is mounted 0.7 m above ground, which creates a dead zone around the robot of approximately 4-m radius. A Hokuyo UTM-30LX scanning laser sensor is installed at 0.25 m for obstacle detection in the dead zone. Within the body of the robot are 4 Mac Mini machines, each with 2.3-GHz quad-core processors and 8-GB memory. The computers run software modules from researchers at different institutions. These different software modules are integrated through the RFRAME framework developed at General Dynamics. RFRAME is a transport agnostic middleware, supporting multiple simultaneous protocols (e.g., Joint Architecture for Unmanned Systems, Robot Operating System, and Neutral Message Language). By abstracting and optimizing differences between environments, RFRAME allows researchers to work in their preferred software environment but as part of an integrated system. The RFRAME system, along with low-level planning and platform control, runs on 1 of the 4 computers.

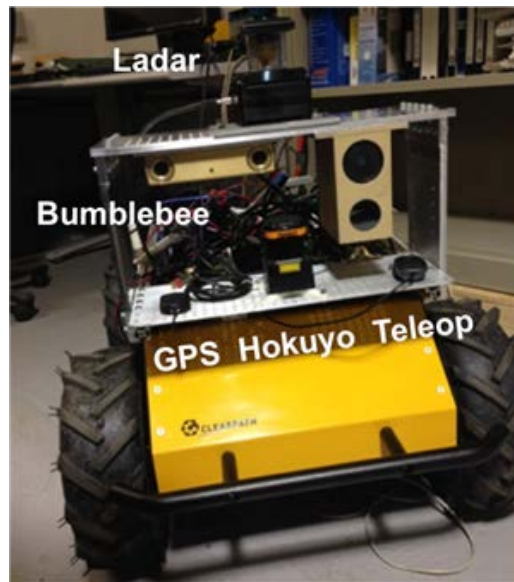


Fig. 2 Husky UGV configuration for semantic perception and navigation IRA

### 4.1.1 Common World Model

The intelligence architecture is built around a Common World Model (CWM) (Dean 2013). This world model combines data that is metric (e.g., sensor data and aggregates), and semantic (e.g., class descriptions and instances), with the robot's

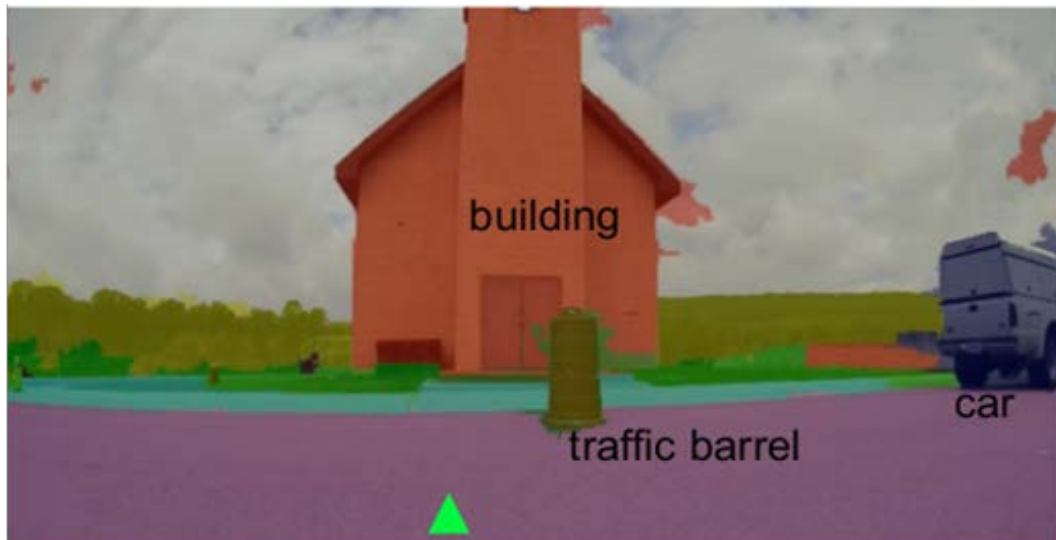


self-knowledge (e.g., position, mission status, goal). The world model is an intelligent data store and not just a database. Internally, the world model knows how the various data sources inter-relate, and, when appropriate, propagates changes between the metric, semantic, and self-levels. At the metric level, CWM efficiently represents and updates sensor data taken from a robot's environment. At the semantic level, objects represent symbolic information, enabling the abstract reasoning needed for intelligent behavior. Finally, self-information contains data relative to the robot itself. Tracking self-knowledge such as current capability, component status, and task execution states, enables the robot to reason, and to adapt its performance.

#### 4.1.2 Perception Method

The detection of different types of objects requires different perceptual algorithms. First, a semantic classifier is used to classify regions of camera images (Munoz 2013). Each pixel of the 2-dimensional (2-D) image is labeled as being one of several types: building, traffic barrel, car, fire hydrant, grass, tree, sky, asphalt, concrete, or unknown. This semantic labeling of objects was tested, and found to be successful, in an earlier IRA (IRA3).

Figure 3 shows a 2-D image fused with 3-D LADAR data to create colored, semantically labeled 3-D point clouds, based on which an object label is chosen (Oh et al. 2015). Such fusing, applied to a traffic barrel and fire hydrant, is shown in Fig. 4.



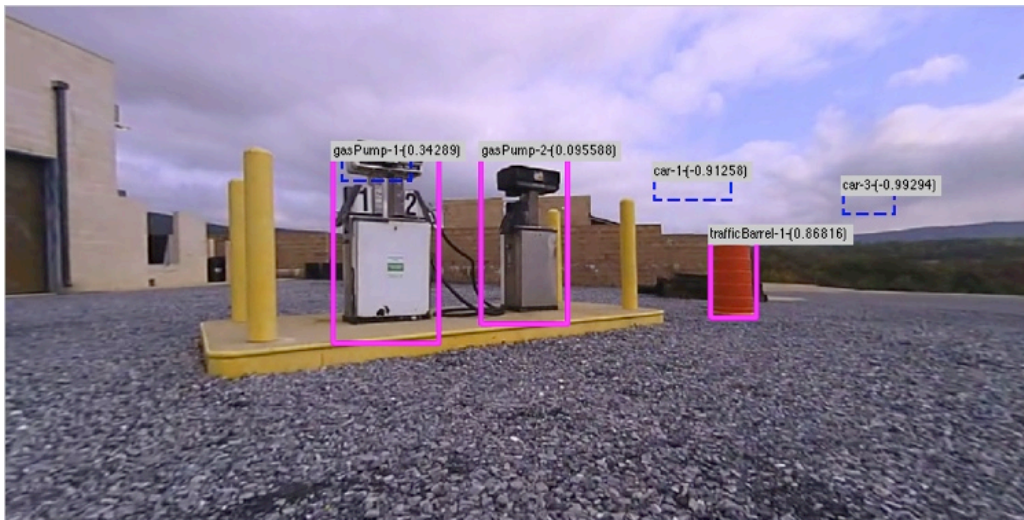
**Fig. 3** A front view of a building, with pixels colored according to semantic label. The building, car, and traffic barrel are labeled with text, and the green triangle shows the position and orientation of the robot.

I.



**Fig. 4** Semantic object detection, with actual objects on the left, and the labeled points of the 3-D point cloud on the right

For detecting objects with distinctive shape features, like the gas pumps shown in Fig. 5, an Active Deformable Part Models (ADPM) method is used (Zhu et al. 2014). This detector was used for the gas pumps and was used in combination with the semantic classifier for traffic barrels, cars, and fire hydrants. Figure 5 shows successful detections of gas pumps and a traffic barrel at the gas station: the numbers are the detection score, the blue dash boxes are false positive boxes in the detection stage, and the red boxes are final detections, which passed the verification stage.

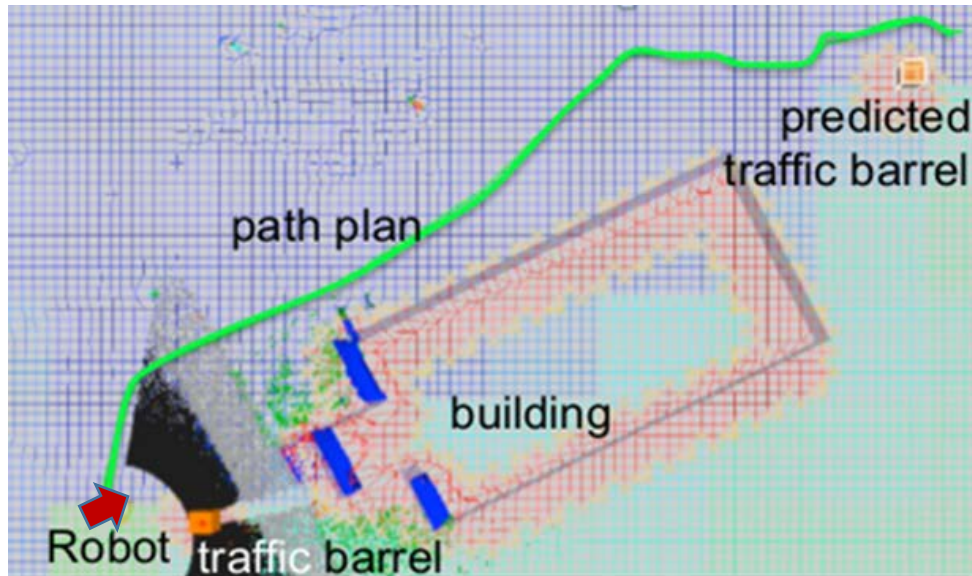


**Fig. 5** ADPM object detection finding gas pumps and a traffic barrel (outlined in pink)

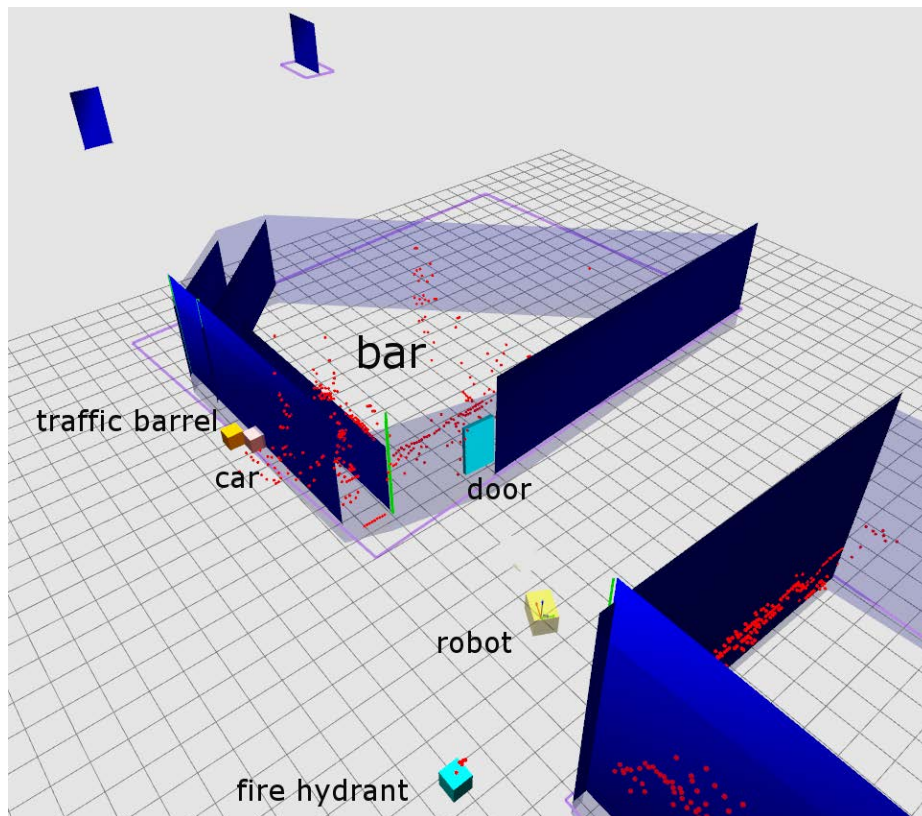
### 4.1.3 Navigation Method

Navigation begins with a command issued to the robot through the HRI interface. This command, called a Tactical Behavior Specification (TBS), is in a structured language that is used for communication among software modules within the intelligence architecture. The TBS language supports a rich set of constraints that leverage spatial relationships among objects in an environment. As an example, consider the command “stay left of the building; navigate to a traffic barrel that is behind the building.” The robot searches the world model for a building in front of it, predicts parts of the building it cannot observe, predicts a position for the traffic barrel behind the building, and plans a path to that goal.

Figure 6 shows the robot’s world model based on the labeled image in Fig. 3. The front walls (in blue) were perceived, as was the traffic barrel in front of the building. The grey walls, and the traffic barrel in back of the building, are predicted objects. In this example, the command includes 2 landmarks, a building and traffic barrel, but the robot’s current world model contains only a set of walls and a predicted building. This inconsistency causes low grounding confidence, which, in turn, enables geometric spatial reasoning. Based on the context in the command, a traffic barrel must be behind the building, so an object is hypothesized behind the building. Now, the world model includes a building and a traffic barrel, both predicted. After symbol grounding is done with sufficiently high confidence, the robot computes a navigation cost map that best satisfies the action constraint to “stay left of the building,” and plans a path accordingly. The representation of the world model chosen in Fig. 6 was chosen because it was easy to interpret. Perception and prediction of buildings during the complete runs was generally not as accurate or complete as in Fig. 6. Consider the representation of the world model shown in Fig. 7.



**Fig. 6** Features in the world model produced by the semantically labeled image in Fig. 3. The blue walls are observed, and the grey walls are predicted. The robot is represented as a red arrow, and the planned path as a green curve leading to the predicted traffic barrel.



**Fig. 7** The rear of the bar, from the robot's perspective, at the completion of run 15. Perceived walls are dark blue, and predicted walls are grey. Also shown is a fire hydrant and the misperceptions of a traffic barrel and car.



Figure 7 shows the model of the world at the end of run 15, when the robot navigated from in front of the gas station to a position behind the bar. The porch of the bar has been represented as an inner and outer wall. A door has been perceived against a predicted wall, and a car and traffic barrel have been misperceived as being next to the bar. Despite this confused view of the world, the robot did complete the run, getting to the correct position and detecting a pedestrian exiting through the door it was facing. The system appears to be relatively robust to the type of misperception images shown here, as long as the misperception is not an object being used as a landmark or goal.

An assessment of performance of the robot in preliminary semantic navigation and perception experiments, and in door and pedestrian detection, is in Lennon et al. 2015b.

## 4.2 IRA: Search and Observe Doorways

---

A different detector was used for detecting doors. In this instance, the search was sped up by the knowledge that a door can only be located on the vertical surface of a building. Thus the door detection algorithm first detects facades, using input from the semantic classifier, and then searches for doors on those facades. An example of the results of door detection is shown in Fig. 8. An examination of perception in the preliminary experiments is contained in Lennon 2015b.



**Fig. 8** Doors detected on the church

There was also perception software running on the system for detecting pedestrians, but that is discussed later in the section on pedestrian detection. Semantic objects, such as buildings, doors, and pedestrians, detected through these perception approaches are added to the robot's world model and are updated as the robot's viewpoint changes over time. All mission commands and planning are interpreted according to the robot's model of the world, which we now describe.

#### **4.2.1 Search**

The search action positions and orients the robot, relative to an object of interest to the human teammate. For example, with the command to “screen the back of the building”, the detected building in the world model is the goal and the robot would reorient toward the center of the building to complete the navigate action. Once this orientation was achieved, the mission planner directs the search action to begin and provides the type of object to search for (a door, in this assessment). The door detection algorithm is always running as part of the perception system, so doors in the scene might already be registered in the world model. In case they are not, the action provides a fixed amount of time for the door detection algorithm to report new detections. After this time expires, the search action will report the number of doors that it found within a configurable field of view. The intention is to have the human teammate choose an object from among the multiple options that would be displayed in the HRI (i.e., all doors found within the allotted time). Once the human teammate selects an object, the robot would then reorient toward that object to begin the observe action. This interaction was not tested as part of these preliminary experiments. Instead, the robot was programmed to orient toward the closest door to its current heading vector. Once this orientation was complete, the search action sent a message to the mission planner, and the mission planner directed the observe action to begin.

#### **4.2.2 Observe**

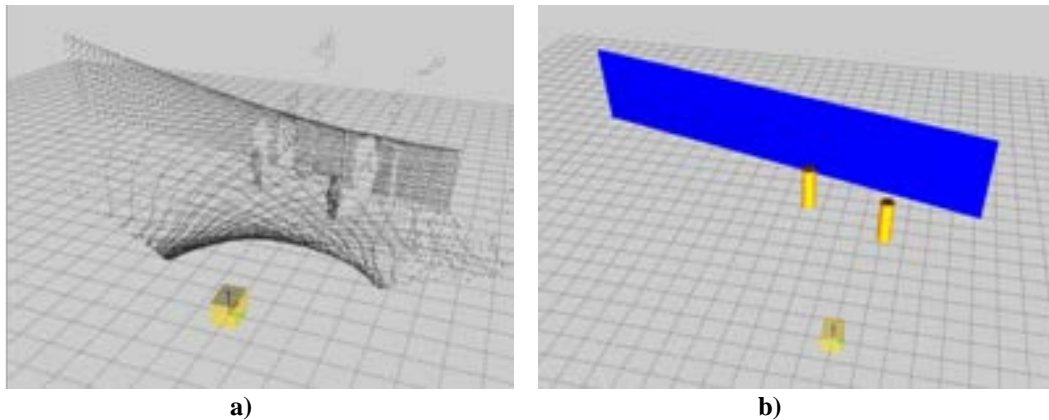
The observe action registers pedestrian detections and reports them to the world model. This action assumes that a previous action has positioned and oriented the robot relative to the object that is being observed. When the mission planner directs the observe action to start, the action begins listening to the output from the pedestrian detector that is already sending pedestrian detection messages. Pedestrian detection messages contain pixel locations for a box that encapsulates the individual parts of the detected person (Yang and Ramanan 2011), and LADAR points within the box are clustered together (Rusu and Cousins 2011). If there are no previous pedestrian tracks of the same shape close by, a new “person” object is added to the world

model. Otherwise, if there is a track nearby that matches in shape, that track is updated. For this assessment, the robot continued to observe in this state until the system was shut down.

In the observe action example shown in Fig. 9, two people exited from the middle and right doors on the back of a building and stood stationary for approximately 5 s, allowing the pedestrian detection algorithm to publish detection boxes and correlate LADAR points in 3-D. They then walked adjacent to the back of the building until they were out of the LADAR's field-of-view. As shown in Fig. 9, the lighting during this portion of the assessment provided challenges to the pedestrian detection. Figures 10a and 10b show the pedestrians in the world model as point clouds on the metric level (Fig. 10a) and as semantic objects (Fig. 10b).



**Fig. 9** Correctly placed detection boxes from the pedestrian detection algorithm



**Fig. 10** These images show representations of pedestrians on the a) metric and b) symbolic level

The sequential execution of the navigate, search, and observe actions constitutes a complete mission.

### 4.3 IRA: Human Robot Interaction Modalities

This effort was a collaboration of RCTA partners at the University of Central Florida, MIT, and ARL, which took place at the CACTF site in the vicinity of the church building. It consists of an assessment in communications, where data were collected using a multimodal interface comprised of speech, gesture, touch, and a visual display to command a robot to perform semantically based tasks. Prior to the data collection, a multimodal user interface (MMI) was used to integrate several research products into a usable means of bi-directional communication with a robot (Fig. 11). The robot incorporated RCTA research software and sensors for planning, navigation, and semantic perception and understanding.



**Fig. 11** The tablet and glove used for and HRI interface. The left side of the tablet’s screen displays a view combining an a priori map from OpenStreetMap (open) with objects in its own world model. The right side, from top to bottom, shows a view through the robot’s camera, the command it is executing, and the activity it is trying to perform.

The human commanded the robot using speech to navigate to different goal locations using, for example, the directive to “...navigate quickly to the traffic barrel near the car”. The robot then used its semantic perception to identify traffic barrels and the car, determine which barrel was near the car, and then navigate to the desired goal location (Fig. 12).



**Fig. 12** Speech is used to command the robot to maneuver in the environment



While the robot navigated, the participant used speech or gestures to convey commands to pause, reorient the robot, or abort and reissue a new directive. Multiple vignettes, all involving complex speech directives and simple commands, were used to examine both the usability of the multimodal interface and the expectations of the human with respect to the behavior of the robot as it carried out the command. In one vignette, an ambiguous situation was purposefully presented. Two barrels were placed equidistant from the building, and the robot was commanded to “navigate to the traffic barrel near the building”. When the robot could not identify which barrel was nearest, it would ask the human to disambiguate the command and choose the correct barrel, currently performed by choosing the correct barrel on the visual interface.

The effort was successful in that the independent research results were integrated into a usable interface that performed some level of bi-directional communication with the robot. Observations on usability and participant expectations with respect to the interaction with the robot were obtained (Hill et al. 2015; Barber et al. 2015). Initial results reveal that there are several usability issues that must be addressed related to the display, speech, and gestures. First, the tablet should provide additional drag and drop capabilities, particularly for map functions. Speech commands are currently constrained, so movement toward more natural military communication styles would enhance the usability. Gestures were considered easy to learn but might be fatiguing over time. Expectations regarding robot behaviors were also obtained, with analysis still in progress.

Suggested improvements are planned to be incorporated in the MMI. Information on human expectations of robot performance will be shared with the robot intelligence developers as a basis for improvements to robot behaviors and to the MMI developers for improvements to bi-directional communications between humans and robots.

#### **4.4 End-to-End Scenarios**

---

In addition to the previously described 4 IRAs of Human Robot Interaction Modalities; Semantic Navigation and Perception; Search and Observation of Doorways; and Search and Grasping of Objects in an Indoor Environment, a fifth IRA was conducted to examine the performance when these capabilities are concatenated. This stringing together of mission steps provides a 2-fold benefit: an appreciation of the performance in a more realistic application and the opportunity to discover if any benefits or shortcomings arise in the interaction or hand-off between segments.

#### 4.4.1 Evaluating the System

The complete runs were performed at the CACTF, around the church, bar, and gas station, which are labeled in the overhead view of the CACTF in Fig. 1. The bar has a complicated facade, with doors and windows set back from the street by several feet. It is also surrounded by other buildings, requiring the semantic navigation system to use landmarks or HRI if the robot starts far enough back from the bar to have several buildings in view. The back of the bar is simple, but the doors are of a color different from the red brick of the building. The space between the gas pumps and the building behind it is used as a storage area for metal lockers, concertina wire, and metal barrels, providing a cluttered environment within which perception was difficult. We used a starting position in front of the gas station to create a situation in which HRI was needed to select the appropriate building. The church is a simple building without clutter, except for trash cans in one front corner. It stands apart from other buildings, is made of cinderblocks, and has tall windows with grey wooden shutters, and grey doors. While it is a simple building to distinguish, we expected that door detection might be more difficult with doors and tall shutters of similar color, and both of a color similar to that of the building.

#### 4.4.2 Design for the Experiment

There was time for 17 complete runs. Table 1 summarizes details of the runs, while Fig. 13 shows positions referenced in that table and in the following descriptions. During runs 1–6, the robot started in front of the bar (B1) and was expected to end behind the bar (BR), with the robot facing a door at the back of the bar. These runs were intended to provide a simple mission executed on a complicated building. Once in back of the bar, however, the search was expected to be easy, as the 3 doors were of distinctly different color than the building. In all runs, we sent the pedestrian out through whichever door the robot was oriented toward, expecting only that the robot would choose a door on the correct building. With the system positioned in front of the bar, no landmarks were expected to be needed, so in each run, the TBS was “screen the back of the building.”

**Table 1 End-to-end scenario run details**

Runs	TBS	Start	Goal
1–6	Screen the back of the building	B1	BR
7–9, 14	Screen the back of the building behind the car	B2 (run 7 only), B3	C1
10–13	Screen the front of the building	C2	C1
15–17	Screen the right of the building that is left of the gas pump	G	BR



**Fig. 13** Test site positions referenced in Table 1. The position of the HMMWV in runs 7–9 and 14 is denoted as H.

For runs 7–9 and 14, the robot started from a position near the bar (B2 for run 7, and B3 for the rest) and was expected to end on the side of the church (C1). For these runs, a landmark was needed to remove ambiguity about which building the robot should go toward, so a High Mobility Multipurpose Wheeled Vehicle (HMMWV) was placed in the street (H) between the bar and the church, and the robot was directed to “screen the back of the building behind the car.” Here, the mission was complicated, but the building at the objective was simple. We thought the robot might have trouble distinguishing between doors and shuttered windows at the church, as both were of similar size and color, so we considered the church to be potentially more challenging for door detection.

In runs 10–13, the robot was placed on the side of the church (C2), and directed to screen that same side (C1) (i.e., “screen the front of the building”). These runs were expected to be straightforward, and easy for the robot, except possibly for the door detection.

For runs 15–17, the robot started in front of the gas station (G) and was expected to go to a position behind the bar (BR). Given the angle of the robot to the bar, the position BR was considered as being on the right of the bar. With several buildings in view, a navigation landmark was needed, so the TBS was “screen the right of the

building that is left of the gas pump”. Even with this command, 2 buildings would have been reasonable choices, and the HRI interface was used to disambiguate the command, directing the robot to the correct building.

#### 4.4.3 Evaluation Criteria

Semantic navigation was evaluated by a human observer, who graded each run on a scale of 0–100, with gradations of 20 (i.e., 0, 20, 40, 60, 80, and 100). The completion score is the subjective assessment of the degree to which the platform accomplished the mission. In Table 2, a score for each run is presented, and for runs scoring less than 100, the reason for the score is listed. In addition, the weather is noted. Runs 1–3 took place on 29 October, during a light rain. As the rain became heavier, the experiment was halted and resumed with run 4 on 30 October.

**Table 2 Scores, weather, and the reason given by the evaluator for scores less than 100**

Run no.	Weather	Score	Reasons for scores less than 100
1	rain	40	Ran into building, may not have seen the wall.
2	rain	100	
3	rain	80	
4	sun	100	Moved to the correct position, but wrong door. Comms failed on HRI
5	sun	100	
6	sun	100	
7	sun	20	Robot went off course and was stopped
8	sun	100	
9	sun	100	
10	sun	100	
11	cloud	40	Software crash during the navigation behavior
12	cloud	60	
13	cloud	100	Software restarted after navigation and before façade detection.
14	cloud	100	
15	cloud	100	
16	cloud	100	
17	cloud	80	In correct position, but pedestrian detection computer had battery failure.

In the Fig. 14 overview, the subjective scores are treated qualitatively to show the variation in performance over the 17 runs. Approximately 65% of the runs were completely successful, achieving a score of 100. Of the remaining 6 runs, there were 2 crashes/restarts of the software (runs 11 and 12) running the platform, one battery failure (run 17), and one communications failure (run 3). During run 1, the robot was stopped because it was going to run into the building. Researchers ascribed this to the network being too slow for updates to the world model to be done in a timely fashion and reduced the quality of the video feed for subsequent runs. The problem did not recur. During run 7, the robot went completely off the

course. Researchers reported that this was because the robot chose to go to a different building than the church. For the subsequent runs from the bar to the church, the robot was moved closer to the church so as to see the church more clearly.

<b>Score</b>	<b>100</b>	<b>80</b>	<b>60</b>	<b>40</b>	<b>20</b>	<b>0</b>
<b>No. runs</b>	11	2	1	2	1	0

**Fig. 14** Number of runs achieving each score

#### **4.4.4 Conclusions**

The end to end runs differed in purpose from the experiment described in Lennon’s research (Lennon 2015b). While that experiment was intended to explore the limits of the system’s ability to reliably understand varying levels of complexity in perception and navigation commands, detect doors, and track pedestrians, the runs presented here were intended to explore the system’s ability to execute sequences of different behaviors. The robot performed a chain of tasks, each of which had been determined by the results of Lennon’s work (Lennon 2015b) to be within its capabilities. Consequently, the navigation commands and perceptual environments presented to the robot were not as complicated as those presented to the system in the assessment (Lennon 2015b), and none of the failures in Table 2 were attributed to failures of semantic navigation, nor to failures in door detection or pedestrian detection. This was not because of improvements in the capability of the robot between the experiments but because it was being given tasks that were known to be within its present capabilities. The system was generally successful in transitioning from one behavior to the next, and we expect that, if it had been given more complicated tasks (as in the assessment [Lennon 2015b]), it would still have transitioned successfully if it had been able to complete those tasks. We intend to test this in the future in the context of evaluating improvements in HRI, which we expect would allow the robot to recover from a failure in one task (e.g., navigation) and continue on with the rest of the screening mission.

### **4.5 IRA: Indoor Search and Grasp**

#### **4.5.1 Mission and Site**

One notional mission, the “Get Object from Inside Building” scenario, is an autonomous indoor search and grasp activity that draws upon research subtasks from the RCTA Annual Program Plan in perception (locating objects within a scene and relative to a manipulator arm), intelligence (room search and approach of objects), and dexterous manipulation and unique mobility (DMUM)(control of mobile manipulators). This section reports the experimental conduct and

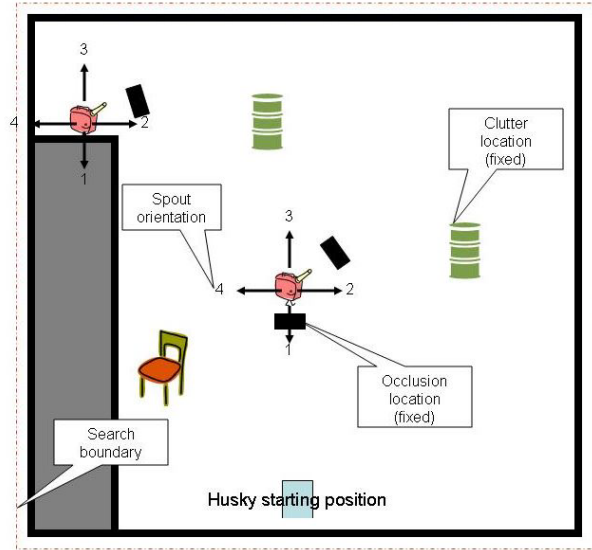
summarizes the performance results for the indoor search and grasp capability. Further detail on the experiment and results can be found in a conference paper (Bodt et al. 2015). That paper is excerpted in this section for the purpose of providing an overview of this IRA activity.

We investigated the “Indoor Search and Grasp” capability to establish a baseline of performance for this first instantiation of the integrated technologies while providing an experimental record to support detailed failure analysis to assist researchers in making system improvements. The Indoor Search and Grasp IRA represents the first time the component technologies were integrated; consequently, IRA conditions as to the notional mission and the challenge of the relevant environment are devised to span an anticipated easy-to-hard space of run conditions where the most can be learned from the assessment. The integrated system searches the room for a specific object, identifies the object, positions the robot so the object can be grabbed by the arm/end effector, and then grasps, lifts, and stows the object to return to start. Observational data on mission outcome along with automatic data collection on the time spent in each subtask of the mission whole were recorded to support the assessment.

All experimentation was performed at the CACTF and the High Bay area of the adjacent ARL Robotics Research Facility. The exploratory runs comprising the bulk of the testing were performed in the High Bay area (Fig. 15) with several confirmatory runs following in the CACTF Firehouse. The walls augmented by plywood and posters constituted a natural boundary for the experiment, and the room-mapping software provided a software boundary keeping the robot within the approximate 30- × 30-ft experimental area where no wall was present. A schematic (Fig. 16) shows the sets of nominal locations and orientations for the gas canister and potential clutter within the room.



**Fig. 15 FTIG High Bay experimental area**



**Fig. 16 Schematic of High Bay experimental area**

#### 4.5.2 Physical System

The key features of the physical system consist of the hardware identified within Fig. 17. The Clearpath Robotics Husky serves as the mobile platform.\* Hokuyo LADAR (UTM-30LX-EW) located in the front of the platform provides sensing for obstacle detection and room mapping.† An ASUS Xtion PRO LIVE (RGBD sensor) supports fine-grained 3-D localization of the target object, which is the red gas can also shown in Fig. 17.‡ A Point Grey monocular camera (CMLN-13S2C-CS) captures images used for initial detection of the target object and coarse localization.§ An HDT Global arm (MK2 Family-Semi Custom 7-DOF Arm w/o Hand) reaches to the target object and lifts after the grasp has been made.\*\* Finally, a RobotIQ gripper (2-Finger 85, Adaptive Gripper) executes the grasp of the gas canister.††

\* [www.clearpathrobotics.com/husky/tech-specs/](http://www.clearpathrobotics.com/husky/tech-specs/)

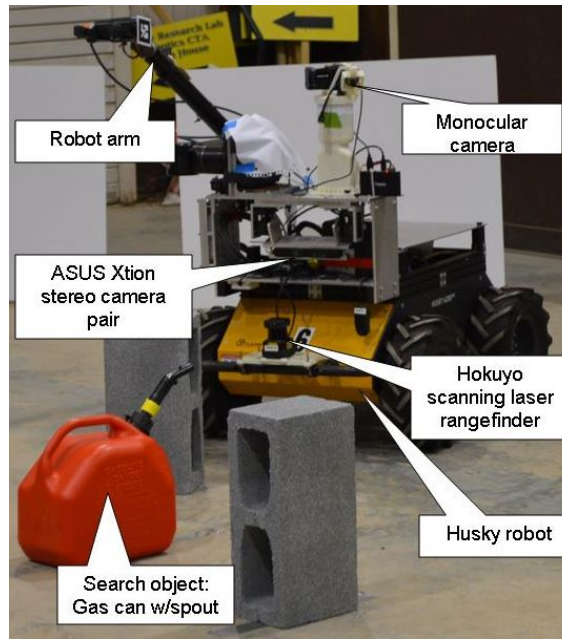
† [www.autonomoustuff.com/hokuyo-utm-30lx-ew.html](http://www.autonomoustuff.com/hokuyo-utm-30lx-ew.html)

‡ [www.asus.com/us/Multimedia/Xtion\\_PRO\\_LIVE/specifications/](http://www.asus.com/us/Multimedia/Xtion_PRO_LIVE/specifications/)

§ [www.ptgrey.com/chameleon-usb-cameras](http://www.ptgrey.com/chameleon-usb-cameras)

\*\* [www.hdtglobal.com/services/robotics/adroit-manipulator-arm/](http://www.hdtglobal.com/services/robotics/adroit-manipulator-arm/)

†† [robotiq.com/products/industrial-robot-gripper/](http://robotiq.com/products/industrial-robot-gripper/)



**Fig. 17 Experimental system components**

### 4.5.3 Experimental Design

A  $2^2 \times 4$  factorial design in 2 blocks was constructed in 4 variables: orientation, occlusion, clutter, and location, with location serving as blocks.

Two locations for placement of the gas can were used: middle of the room and against a wall. They were fundamentally different locations. Placement of the search object (gas can) against a wall would significantly affect the ability of the robot to locate and grab the gas can. Against a wall creates 2 challenges: 1) it makes the depth determination more difficult and 2) it reduces the number of potential grab locations available to the robot.

Four gas can orientations were used and defined by the orientation of the spout as seen from the starting position (Fig. 16). Position 1 is with the spout pointing toward the starting location; position 2 is with the spout pointing to the right perpendicular to the line of sight from the starting location; position 3 is with the spout pointing away from the starting location; position 4 is with the spout pointing to the left perpendicular to the line of sight from the starting location. The orientation was expected to present a challenge to the robot for 2 reasons: 1) the gas can identification requires the spout in sight to determine the can azimuth, and 2) because there is a preferred grab location to the left rear of the can that would obviously change with each can orientation.



Occlusion objects were solid objects placed within 12 inches of the gas can, which would occlude the ASUS view of the gas can if it was in line with the approach taken by the robot. This challenged the robot in 3 regards: 1) in identifying the gas can when it was occluded to some degree, 2) by presenting an obstacle to achieving a suitable grab position, and 3) by presenting an obstacle that would potentially interfere with the planned arm trajectory to grab the can, thereby reducing the number of sufficient grasp locations.

Clutter consisted of barrels and chairs placed within the search area but at least 8 ft away from the gas can so that the clutter would not interfere with the ASUS vision system or the grab location calculations when the robot was in close proximity to the gas can. Clutter objects challenged the robot as an obstacle to be avoided and by making the mapping movements more difficult and time consuming. What path the robot chose to avoid clutter during room search did have the potential to influence the perspective of camera shots. Otherwise, it was thought that the clutter variable (clutter/no clutter) would not influence the final outcome (success/fail).

#### **4.5.4 Metrics**

There were 3 possible outcomes for each run. A success was recorded if the robot identified the gas can, grasped it, and lifted it up. A partial success was recorded if the robot identified the gas can, moved to a viable grab location but failed to grasp the can because it was still too far away or because something potentially interfered with the planned robot arm path. After several repeated attempts with the same result the test director ended the mission. A failure was recorded if the robot failed to see the gas can, or having identified the gas can, failed to move to a good grab position. Each run recorded as a failure was ended at the discretion of the test director when it was clear that no successful grab was likely.

In addition to the 3 run outcomes, a number of response variables were collected when the run was successful or partially successful. For all successful runs, we collected 5 additional variables: 1) mission run time from start to can grab, 2) time taking pictures and processing the data, 3) time planning and moving to map the room, 4) time positioning the robot to a grab position, and 5) time spent grabbing the gas can. For all successful and partially successful runs, we collected the following data: 1) time to first identification (ID) of gas can, and 2) time from first ID to first grab attempt.

#### **4.5.5 Results Commentary**

The highlights of the investigation follow. Of 25 runs performed as part of the principal experiment, the system accomplished 15 (60%) complete successes, 7 (28%) partial successes, and saw only 3 (12%) failures. An excursion in a more

realistic, cluttered environment, saw 7 of 8 runs as complete success and one failure. The orientation of the gas can with respect to the robot does affect the gas can identification. Specifically, the spout is an important feature in making the target determination. Objects near the gas can reduce the number of successful grabs, not because of visual occlusion but because of difficulty in achieving a suitable grab location close enough to the can to grasp it and yet free of interference with the planned arm movement. Room clutter, other than altering the path of the robot during search, did not have an impact on performance. The current configuration of the camera and arm, along with algorithm decisions on how many pictures were taken and frequently created blind spots over the ‘right shoulder’ of the robot, did impact its ability to clearly see the target object when navigating counterclockwise around the room. For successful runs when the gas can was in the middle of the room, the completed run time was on average approximately 3 1/2 min and near the wall approximately 8 min. The longer time for the wall runs was due to the search required; once the gas can was identified, there was no appreciable difference in times associated with positioning for and executing the grasp.

## **5. Capstone Experiment Task-Based Assessments (TBA)**

---

In addition to the previously described IRAs, during the timeframe of the Capstone Experiment, 5 TBAs using various platforms were also conducted: Bracing to Reach and Grasp an Object, Detection and Climbing of Stairs, Leaping over a Span, Dynamically Feasible Motion Planning, and Terrain Aware Motion Planning.

### **5.1 TBA: Self-Anchored Reaching (JPL)**

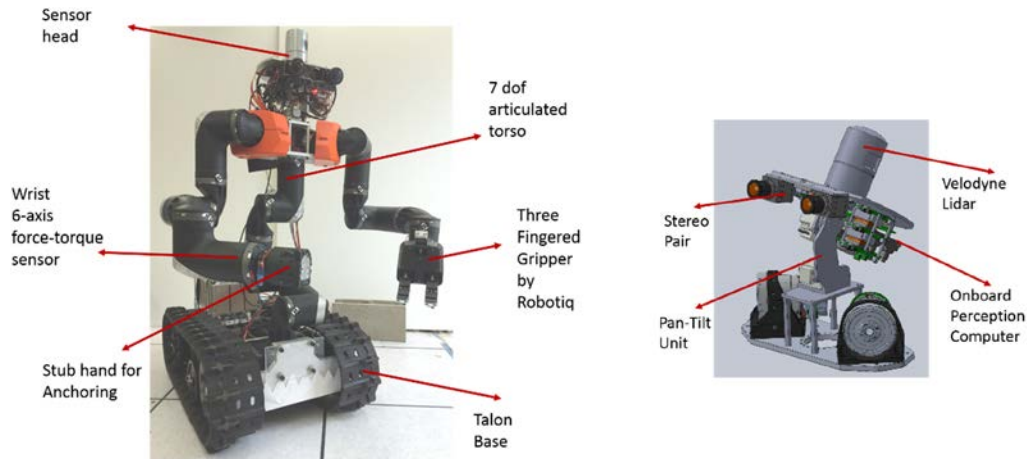
---

#### **5.1.1 Introduction: Overview**

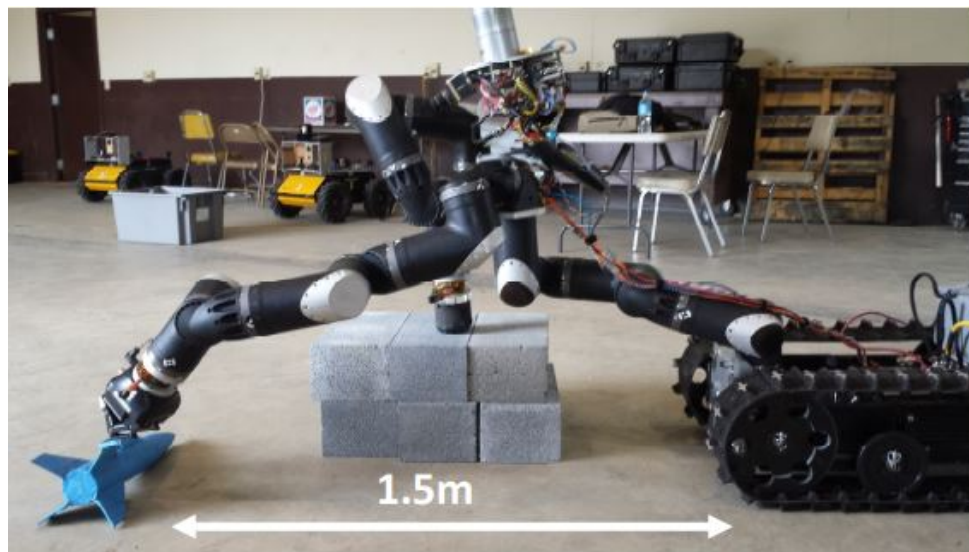
This section highlights results from TBA of Sensor-Based Dexterous Manipulation and High Performance Visual Range and Motion Estimation for Small Platforms. These activities were administered partially at JPL and partially at FTIG.

The capabilities addressed in these assessments will enable robots to perform advanced behaviors in finding objects of interest hidden in hard to reach places. In particular, the pieces relating to an operator specifying (talk) gaze goals for exploration (look) followed by the robot navigating and moving its body into configurations that interact with the world (think, move) are highly relevant to the RCTA vision.

The focus of the Sensor-Based Dexterous Manipulation assessment is to interact with the environment by closing a kinematic chain on the surrogate platform (Fig. 18) by anchoring one hand with the environment and extending the other to reach and grab an object (Fig. 19). The objective is to enable the robot to perform coordinated whole-body movements while anchored to the world at 2 points. Without compliance in control using force-torque sensor measurements, the system would be too stiff and any whole body movement can either disturb the world or inflict self-damage to the robot.



**Fig. 18 The surrogate platform and the 360° field of view sensor head**



**Fig. 19 Extended reaching (~5 ft) in a bracing maneuver**

The objective of the High-Performance Visual Range and Motion Estimation for Small Platforms assessment is to enable the operator to specify end-effector and gaze goals a large distance away ( $>5$  m) from the experimental setup. In the absence

of a motion estimation module from perception, the reference frame in which the goals are specified would be invalidated over time with subsequent robot motions due to visual odometry drift. The objective of maintaining a common reference frame for use within the manipulation task will be achieved by integrating 1) path planning and path following for the base to achieve navigation goals and 2) whole body motion planning and control given end effector and gaze (look-at) goals from an operator.

### 5.1.2 Experimental Design and Task-Based Assessment Criteria and Metrics

At the gap we performed about 8 runs of self-anchored pick and place of an object of interest. Table 3 outlines the task outcomes, execution times and specific failures.

**Table 3 Self-anchored reaching run details**

Run	Execution time	Outcome	Comments
Oct 5th run 1	10 min	Success with 1 grasp failure	Grasp failed first time, had to re-grasp. Fingers hit the ground.
Oct 5th run 2	7 min	Success	...
Oct 5th run 3	9 min	Success	...
Oct 6th extended run1	18 min	Success with 1 grasp failure	Run included removing a gas canister out of view and then picking up an object of interest. The initial grasp of the gas canister was unsuccessful as the grip was not tight enough. A second attempt succeeded but with only a 2-fingered grasp.
Oct 6th run 2 at test range	9 min	Success	...
Oct 6th run 3 at test range	8 min	Success	...
Oct 7th run 1	12 min	Partial success	Initial grasp failed due to stale maps (operator error). The second grasp succeeded.
Oct 7th run 2	7 min	Success	...

The Sensor-Based Dexterous Manipulation task was assessed with 2 metrics: Reachability Gain and Disturbance Rejection

For the High-Performance Visual Range and Motion Estimation for Small Platforms task, the goal was to assess in terms of repeatability of the experiment from different starting positions and validity of motion goals over time which is measured via motion drift. It is anticipated that successful DMUM operation would require motion drift on the order of 10 cm.

### Metric 1: Reachability Gain with Anchoring

The increase in torso range with and without anchoring. This will be measured in terms of how far the center of mass can be outside the base support.

Without bracing: The robot tips over when the center of mass (estimated) is greater than 0.35 m from the center of the base.

With bracing: The robot can extend to edge of the kinematic reachability gain beyond which the planners fail. In practice we were able to extend the center of mass to 0.5 m from the center of the base before the planners did not return solutions. This is a 30% increase in position of the center of mass relative to the center of pressure (Fig. 19).

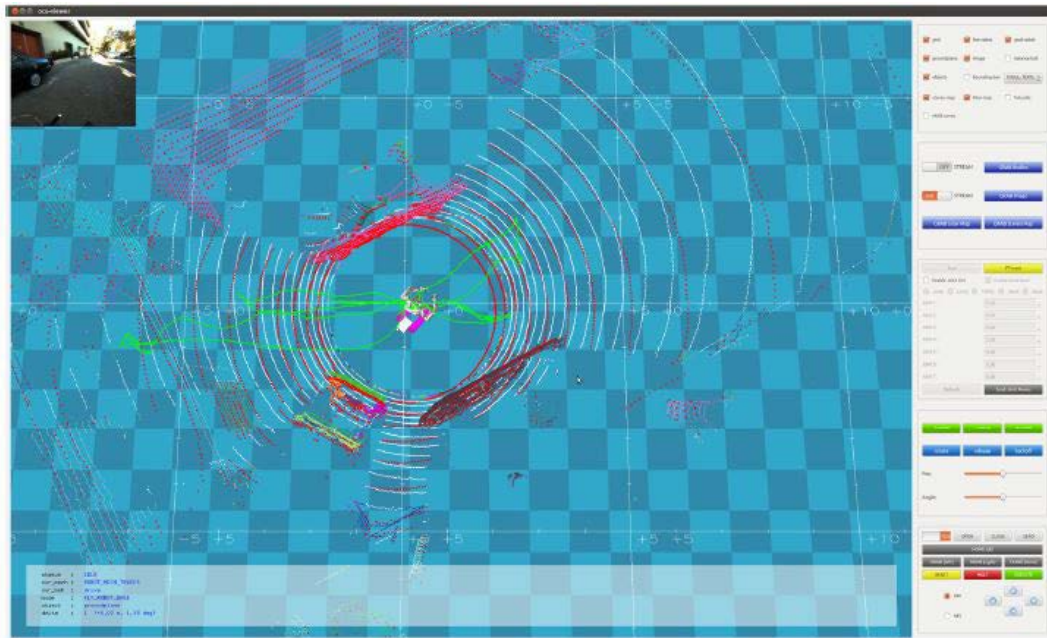
### Metric 2: Setup for Disturbance Rejection Experiments

For assessing disturbance rejection, the robot was started in a bracing position and a select torso joint was actuated to apply a step disturbance.

Perception Metric 1: End-to-End Navigation Error Experiments with LADAR Odometry Navigation error, at the end of plan, of the robot, and commanded goal locations. Repeatability is a strong function of how often the robot can navigate to a fixed goal reliably with state estimation.

### 5.1.3 Results

The mobility tests consisted of 5 separate runs with the robot placed at a known start location, manually driven approximately 5 m away, and commanded to autonomously navigate back to the origin while using only the LADAR pose solution. Figure 20 shows the robot path taken during the 5 mobility test runs, overlaid with LADAR point clouds.



**Fig. 20 Multiple navigation runs with the a goal of {0,0} and different start locations**

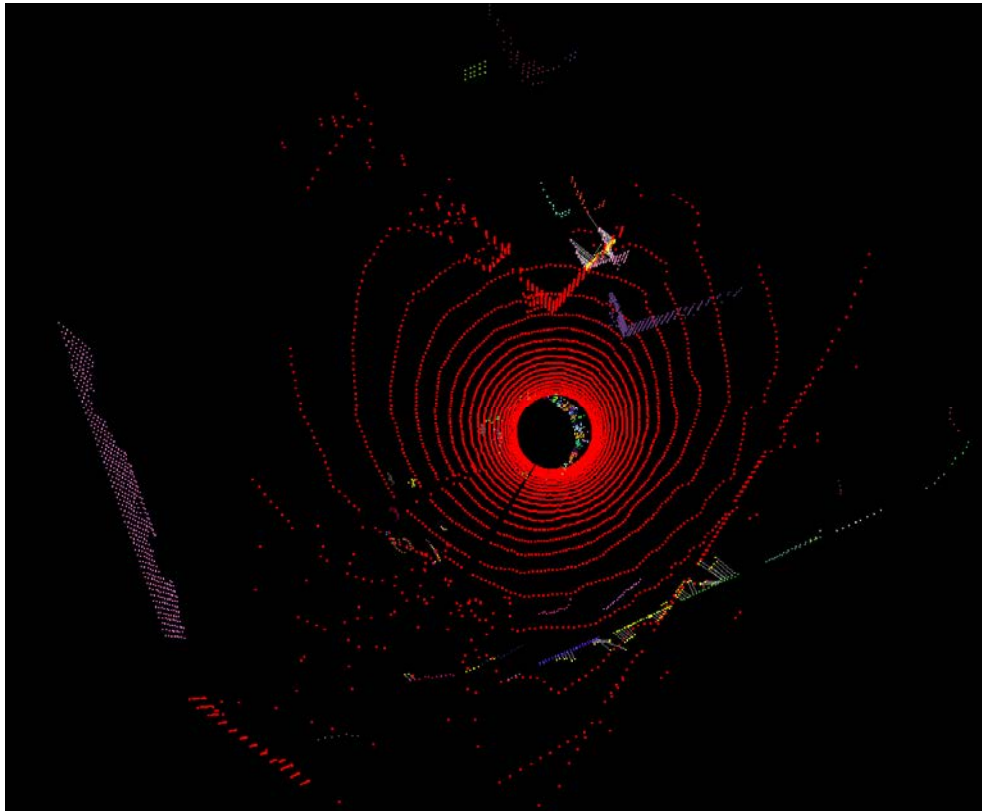
The drift of the navigation solution was measured by the displacement between the origin of the trajectory and the pose of the robot after navigating back to the origin. The 5 runs resulted in the drift shown in Table 4. The paths taken were approximately each 5m long. The average drift across the runs in the table is 10 cm. The threshold for the mobility planner on reaching the goal was 5 cm, and the grid size used in the map for planning was 5 cm as well, so the pose accuracy was likely better than 10 cm. These results capture system-level navigation capabilities since several modules were run to produce these results, including track control, D-star navigation, perception, mobility, manipulation, and pose estimation.

**Table 4 End-to-end navigation error, at the end of plan, of the robot and commanded goal locations**

Run no.	Distance (cm)
1	12
2	10
3	8
4	9
5	13

These tests were conducted with the vision pipeline turned off, that is, without stereo computation and visual odometry (VO) running. When full integration was tested, including stereo and VO in the main perception loop, the overall rate of the perception stack went down from above 10 Hz to about 2 Hz. This caused an

increased drift in the LADAR-based odometry compared with the results shown previously. This is due to the larger displacement of the robot between 2 perception updates (since the perception rate was then slower). This is not due to a fundamental limitation of the approach and can be solved through software engineering. We are in the process of reorganizing the perception code so that LADAR-based odometry runs in its own process and does not compete for central processing unit time with the rest of the perception pipeline. This will allow the same level of accuracy as described in Table 4 while running the full perception algorithms suite. Figure 21 shows an instance of large displacement between 2 consecutive scans due to a lower perception rate. In that case, the perception module had dropped a number of scans between these 2 scans (due to the low update rate) resulting in a large displacement between the 2 scans processed. The white segments indicate the point associations found by the alignment algorithm. As can be seen, these are incorrect due to the initial large displacement. This can be solved by processing incoming scans faster and thus processing more scans with smaller displacements between them. The aim of the current software updates is to allow such faster LADAR processing.



**Fig. 21** Example of large displacement between 2 LADAR scans when running the perception pipeline at low rate. The reference scan is shown in red, the current scan that will be aligned to the reference scan is colored by segmentation. As can be seen, the main structures in the scene are off by about  $45^\circ$  between the 2 scans.



#### **5.1.4 Conclusion**

The TBA tested the integrated semi-autonomous capabilities for grasping a target object with bracing using the surrogate mobile manipulation robotic platform. The tests were broken into 1) the repeatability of the manipulation tasks, 2) the robustness of bracing behaviors as a function of force and current loads that build in the joints segments, and 3) the end-to-end navigation accuracy of system. The current maturity of the technologies evaluated in the TBA is not yet at a level that would enable fully autonomous mobile systems. A human operator is still required to make high-level decisions and judgments for successful task completion. However, results and performance look promising, and the semi-autonomous grasping implementation assessed here may be an attractive alternative to complete teleoperation of a robotic system. Semiautonomous robotic systems would allow the operator to expend more of his energy and focus on his safety and surroundings than traditional teleoperation.

### **5.2 TBA: Stair-Climbing with XRHex**

---

#### **5.2.1 Introduction: Autonomous Stair Climbing Overview**

This report highlights results from a TBA that was administered at FTIG. The TBA evaluated performance of XRHex, the robot hexapod, as it carried out a semi-autonomous stair climbing behavior.

This TBA was largely used as a test of a new stair detector algorithm utilizing an RGBD camera and a switching behavior written to begin a previously implemented stair climbing gait upon stair detection. The behavior is considered to be semi-autonomous because it still requires an operator to position the robot in an acceptable starting point with respect to a flight of stairs for the autonomous behavior to trigger. Successful trials during this TBA would encourage further development of the switching behavior to a point where XRHex is capable of positioning itself in the appropriate location to begin stair climbing regardless of the range at which the stair case is detected (effectively enlarging the basin of attraction for this behavior).

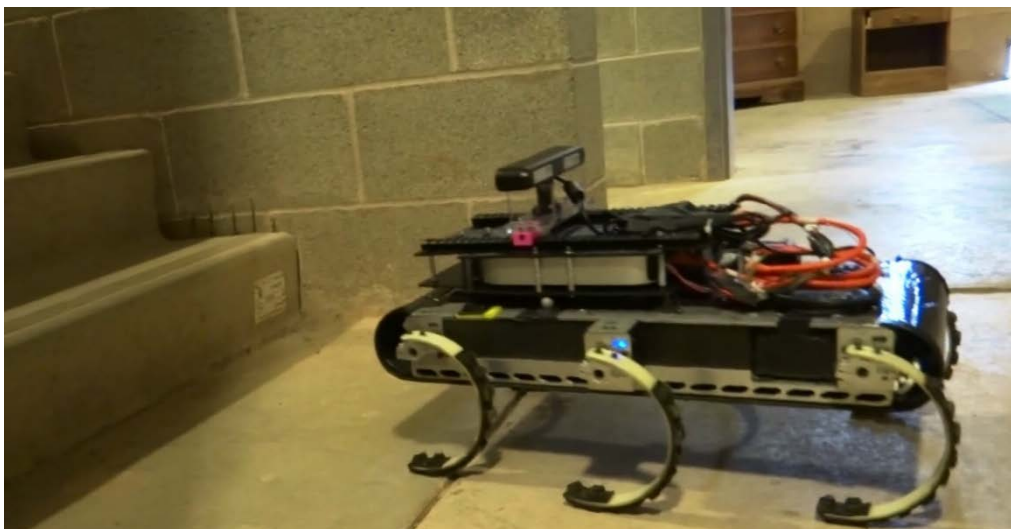
#### **5.2.2 Experimental Protocol**

Experiments were conducted indoors at the hotel and police station. Some initial trials were attempted outside, but the sunlight washed out the depth sensor to the point where stair detection would not be possible. A trial consisted of enabling the stair detection behavior and then having the operator drive XRHex. To test the robustness of the stair detector, XRHex was driven around a full loop of each floor, including going into at least one room, before approaching any flight of stairs to



ensure that we did not have a high rate of detecting false positives. After finishing the initial exploration, the operator drove the robot to a suitable starting position (roughly 50 cm from the base of a staircase with the robot's heading within  $\pm 7^\circ$  of the axis of the stairs). From there, the switching behavior activated the stair climbing gait, XRHex climbed to the top of the stairs, detected when it reached a landing, and then re-enabled operator control while looking for more stairs.

The physical setup of the robot can be seen in Fig. 22, and the mass of the robot and sensor payload are listed in Table 5. The payload accounts for roughly 20% of the total mass of the system in this configuration. For all successful trials, the payload was shifted as far forward as possible on the robot's rail mounts. If the payload was located further back, the robot would occasionally pitch backwards and fall while attempting to climb.



**Fig. 22** Experimental setup. An RGBD camera and Mac Mini are mounted on XRHex. Off screen, an operator is holding a joystick used to drive the robot when away from stairs.

**Table 5** xRhex Experimental setup mass measurements

Component	Mass
Robot + 1 battery	9.057 kg
Mac mini (with mount)	1.897 kg
RGBD camera (with mount)	.325 kg
IMU	.052 kg
Total mass	11.327 kg

Note: IMU = Inertial Measurement Unit

The stairs in the 2 test buildings (and across the rest of the test site) were nearly identical, differing only in the number of stairs per flight and shape of the landing between flights. In the hotel, there was an L-shaped landing between flights, and in the police station the single landing was a narrow rectangle. The stairs themselves were made of smooth poured concrete, had rounded noses, and were coated with a

significant amount of dust. All of these factors significantly reduced traction, making stair climbing somewhat difficult. The rise and run of each step can be seen in Table 6. The “ground truth” values were taken with a measuring tape, and all other values were measured by the RGBD camera. The camera measurements were not used for these experiments but they were taken to get calibration data for the sensor.

**Table 6 Stair measurements and robot measurements taken from depth camera**

<b>Measurement method</b>	<b>Rise (cm)</b>	<b>Run (cm)</b>
Ground truth	18.0	28.0
Robot sitting, 1.15 m from 1st step	16.2	26.2
Robot standing, 1.15 m from 1st step	15.5	26.0
Robot sitting, 0.5 m from 1st step	15.5	26.8
Robot standing, 0.5 m from 1st step	16.0	27.0

### 5.2.3 Results and Analysis

Six trials were conducted for the assessment: 3 at the hotel and 3 at the police station. These trials showed that the stair detection algorithm and the behavior as a whole are quite reliable. A summary of the tests can be found in Table 7. The 6 trials took slightly under an hour altogether, during which time the robot saw a variety of furniture including desks, chairs, beds, and tables. None of these objects triggered a false positive stair detection. In an attempt to force the detector to fail after all trials were completed, a robot handler held the robot on its side near a set of vertical bars for a jail cell in the police station. This did cause a false positive, but it only triggered when the robot was held at an angle such that the gaps between the cell bars could not be seen. This implies that parallel bars (like for a sewer grate) may cause trouble for the detector, but the test site did not have any other similar features to test on.

**Table 7 xRhex stair-climbing experimental results**

<b>Trial</b>	<b>Stairs per flight</b>	<b>Flights per floor</b>	<b>Total stairs climbed</b>	<b>Stair slips</b>	<b>Trial duration (min)</b>	<b>Notes</b>
Hotel 1	7	3	63	8	12	Transition to stairs failed once, behavior automatically reacquired stairs and was successful on second attempt
Hotel 2	7	3	63	8	12.5	Robot scraped along wall for one flight of stairs due to poor positioning (operator error), behavior successful
Hotel 3	7	3	63	4	12	Transition out of stair climbing stalled once, requiring operator to manually resume walking phase
Police 1	10	2	20	2	6	One stair detection failure due to sensor timing after making a sharp turn, one transition failure where RHex nearly walked off open edge of stairwell
Police 2	10	2	20	0	4	Route for this trial avoided sharp turn, no errors
Police 3	10	2	20	4	9	Same route taken as in Police 1, stair detection successful

Apart from false positives, the stair detector did fail one time when a set of stairs was located immediately after a sharp turn. This was likely due to the sampling time of the sensor. If the robot sampled during the turn, it is possible that the camera was too close to the stairs to allow for acquisition by the next sampling time. In total, 33 individual flights of stairs were climbed with only one detection failure. Additionally, when the robot was repositioned after the detection failure, the stair detector was able to acquire the stairwell.

The stair climbing behavior also proved to be robust, despite the significant weight of the payload and the condition of the stairs. In normal operation, the robot transitions from walking to stair climbing by finding the first stair with either of its front legs. The single leg should catch on the edge of the first step and then help align the robot's body with the stairs. This part of the behavior did not work as intended because a single leg could not support the combined weight of the robot and payload causing it to slip off of the step (assisted by the low friction of the steps themselves). Despite this, the transitions were still successful because when the legs began to move in pairs, they were strong enough to lift the robot onto the first step. While on the stairs, the robot occasionally slipped as a result of poor leg positioning. This resulted in the robot falling back by one step, catching itself, and then continuing to climb. There was one instance where the robot slipped off the stairs completely, but this was not during one of the recorded trials.

#### **5.2.4 Conclusion**

This TBA demonstrated that the stair detection algorithm and switching behavior are robust enough to be developed further. The main failures of the assessment dealt with slipping on stairs (both as a result of the stair characteristics and a gait untuned for the specific stairs used) and the transition to the first step. Tuning (possibly automated) of the stair climbing parameters for the specific set of stairs used in the demonstration would likely further improve performance; however, the goal was to demonstrate the generality of the stair climbing behavior to any set of stairs. The trials provided good insight into areas that could be improved and have encouraged us to explore implementing an adaptable stair climbing gait, as opposed to the current one that uses predetermined set points. This adaptable behavior will likely need to use some information about the stairs themselves, such as the rise and run, which can come from the sensor, as shown in Table 6. As we improve the stair climbing gait itself, we will be able to get closer to a platform that allows for autonomous multi-floor building exploration.

### **5.3 TBA: Gap-Crossing with Canid Quadruped**

---

#### **5.3.1 Introduction**

This report highlights results from the 2014 Capstone TBA that was administered at FTIG. The TBA evaluated the performance of the quadrupedal robot Canid performing gap crossing maneuvers. This Capstone TBA represented a continuation of the July 2014 TBA that began investigating outdoor gap crossing. While the July TBA investigated the role of Canid's rear legs in forward leaping, the Capstone TBA investigated the sensitivity of forward leaping behavior to the elevation difference and compactness of terrain.

During the 2014 Capstone TBA, Canid was recorded leaping from a Pelican case onto the bank building ledge at the Fort Indiantown Gap, and—for the first time in a natural outdoor environment—at a drainage ditch nearby as pictured earlier. Canid performed well in leaping onto the bank ledge; however, it was not successful in fully crossing the drainage ditch. Valuable lessons were learned from the failures and will help to improve future performance.

#### **5.3.2 Experimental Protocol**

A series of 14 experiments were completed in which the Canid robot (Pusey et al. 2013) leapt from a Pelican case to a concrete ledge in front of the bank building at FTIG. This was only the second time Canid has been tested in an outdoor environment, the first being the previous TBA in July 2014. While the distance between the jumping and landing platform was varied at the previous TBA in July,

at the 2014 Capstone TBA the elevation difference was varied so as to examine the sensitivity in Canid's open-loop leaping behavior to terrain disturbances. Pictures of the test setup are shown in Fig. 23. Details regarding the leaping testing procedure are similar as provided in a paper presented at the International Symposium on Experimental Robotics (Duperret et al. 2014).



**Fig. 23 Canid leaping a span**

Canid leapt from the Pelican case to the ledge over the course of 14 runs. The height of the pelican case compared with the ledge was varied over the course of the runs to investigate Canid's sensitivity to leaping conditions.

Canid was also taken to a nearby drainage ditch (Fig. 24) for additional gap-crossing experiments with permission of the on-duty range officer. This test was intended to investigate the effects of loose terrain on Canid's leaping behavior, in contrast with the rigid structure of the Pelican case and concrete Bank landing. Until this point, Canid had never been tested on natural outdoor terrain such as dirt or grass.



**Fig. 24 Canid leap at drainage ditch**

### 5.3.3 Results and Analysis

The results for the 12 Bank runs are shown in Table 8. The measured distances were done by hand and suffer from the associated measurement inaccuracies as compared with motion capture data that are usually collected. Additionally, the placement of Canid on the starting Pelican case varied between runs, which likely accounts for the majority of the failure cases not related to the malfunctioning power board or leg bearing failure.

**Table 8 Gap crossing data detailing Canid leaps from a Pelican case to the bank ledge**

<b>Trial no.</b>	<b>Initial to landing platform elevation (cm)</b>	<b>Distance from platform to platform (cm)</b>	<b>Crossed or not (yes/no)</b>	<b>Notes</b>
1	+4 cm	42	Yes	...
2	+4 cm	42	No	Failure likely due to incorrect placement on case.
3	+4 cm	42	No	Failure likely due to incorrect placement on case.
4	+4 cm	42	Yes	
5	+4 cm	42	No	Hit front ridge Rear legs didn't kick. Later this was attributed to a malfunctioning power management board.
6	+4 cm	42	No	Rear legs didn't kick. Later this was attributed to a malfunctioning power management board.
7	+4 cm	42	No	New batteries, replaced power management board.
8	+4 cm	42	Yes	
9	+4 cm	42	Yes	
10	+4 cm	42	No	Realized later that a bearing had broken in Canid's rear left leg in trial 10 failure that likely accounted for this failure.
11	-6.5	35	No	Performed jump even though rear left leg crank fell off.
12	-6.5	35	Yes	Performed jump even though rear leg caught.
13	-6.5	35	Yes	
14	-6.5	35	Yes	...

Not including the instances of incorrect placement (trials 2 and 3) and issues with power electronics (trials 6 and 7), Canid was able to complete 4 out of 6 leaps onto a higher ledge. Likewise, not accounting for a mechanical failure related to a previous crash, Canid was able to complete 3 out of 3 leaps onto a lower ledge. This indicates that Canid is not overly sensitive to minor variations in the height of its landing zone and suggests a degree of stability in its leaping. Like the previous



TBA, while Canid suffered multiple leaping failures it never suffered catastrophically, and the required repairs to incorrect landings on concrete were minor, showcasing its mechanical robustness.

After the tests at the range's Bank, Canid was brought to a drainage ditch (with the permission of the on-duty range officer) and attempted to leap across it 4 times with varying terrain conditions. To vary the terrain, Canid was run separately in areas where it was very grassy, where it was primarily dirt, and where there was a mixture of grass and dirt. None of the leaps were successful; however, high-speed video analysis indicated that all of the failures were due to Canid's legs slipping on the loose terrain. The right panel of Fig. 25 shows an example of where Canid's rear right leg lost traction and kicked up dirt and grass instead of propelling the body forwards. As Canid was being run with a "cleated" leg design featuring bolts jutting from its legs to increase ground friction, it is unlikely that these terrain failures resulted from an inherently inadequate coefficient of friction with the ground. It is more likely that there was insufficient normal force for these bolts to catch or compact the loose terrain so as to push off in a successful leap.



**Fig. 25** Canid after failing to cross the drainage ditch (left) and the ground scuffmarks from insufficient traction in Canid's rear legs while jumping (right)

The results of Canid's first foray into realistic, loose outdoor terrain suggest that if Canid is to operate effectively outdoors it must be able to generate varying degrees of normal forces to account for a variety of terrains. Currently Canid possesses single degree of freedom legs and is only able to vary its toe trajectory through mechanical changes to its 4-bar linkages, making on-demand changes to the leg normal forces difficult. One possible solution would be to add an extra degree of leg freedom as to be able to separately control normal force and forwards force throughout the leap. Coupled with a transparent transmission, this could allow for Canid to "feel" ground slippage and adjust its downward leg force as needed to accommodate looser terrain.

### **5.3.4 Conclusion**

The TBA tested the gap crossing ability of the Canid platform in an outdoor environment. A total of 14 open-loop runs were conducted in which the elevation between Canid's leaping and landing platform was varied. When discounting failures unrelated to the difference in heights, Canid was able to successfully leap the majority of the time indicating a lower sensitivity to landing height conditions than previously thought. Canid was also run in dirt and grass for the first time when it attempted to leap across a drainage ditch. It was unsuccessful during these runs, likely due to an inability of its rear legs to generate sufficient normal forces to gain traction in the loose dirt and grass. This motivates the investigation into using 2-degree-of-freedom legs for future testing.

## **5.4 TBA: Efficient Motion with Dynamic and Power Models**

---

### **5.4.1 Description of Capability**

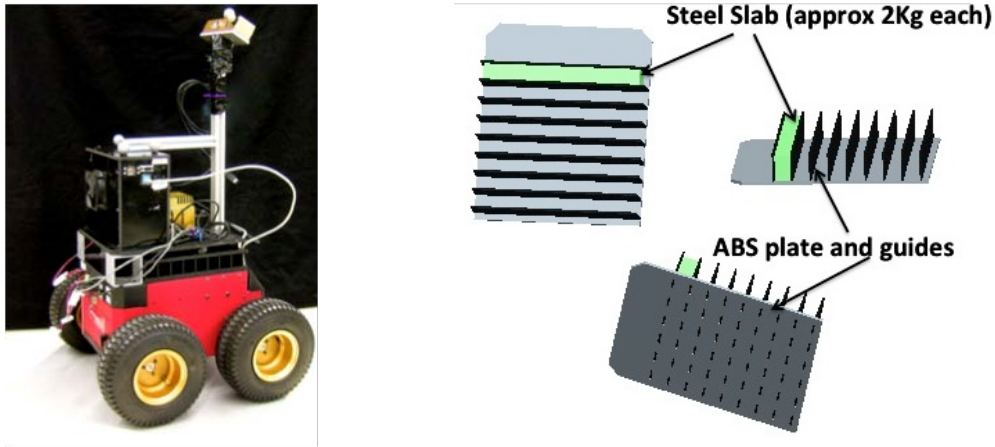
A skid-steered vehicle can be either tracked, legged (e.g., XRhex type), or wheeled and is characterized by 2 features. First, the vehicle steering depends on controlling the relative velocities of the left and right side tracks, legs, or wheels. Second, all joints remain parallel to the longitudinal axis of the vehicle and vehicle turning requires slippage of the tracks, legs, or wheels. However, as with any real platform, skid-steered vehicles have some important limitations. These platforms must slip and/or skid to turn, which makes them less predictable than, for example, differentially driven vehicles. Also, while performing sharp turns, the required motor torques increase significantly when compared with straight-line motion, which can lead to actuator saturation and result in degraded performance.

The primary focus of this assessment was to evaluate terrain and payload dependent slip (Seegmiller et al. 2013) and dynamic and power models for skid-steered vehicles and show their application in energy efficient motion planning (Gupta 2014; Gupta et al. 2015; Ordonez et al. 2015). As part of the assessment, it was shown that when these models are ignored and traditional minimum distance planning is performed, it is possible to develop trajectories that violate the torque constraints of the actuators. For example, this may occur due to high friction between the running gear and surface when making a sharp turn. These trajectories can lead to vehicle stall or poor tracking of the vehicle commands.

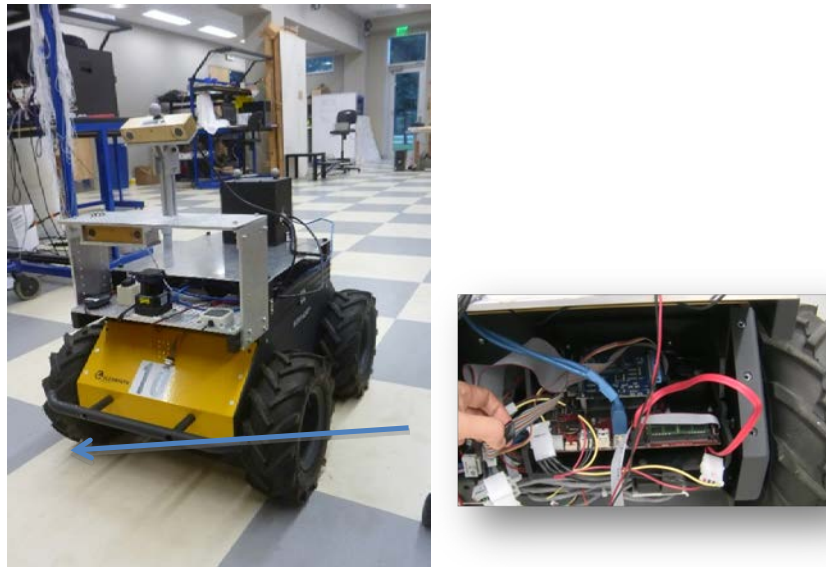
The assessed methodology performs online adaptation of vehicle models by combining detailed slip and terramechanics-based dynamic models of wheel terrain interaction with online learning via an efficient neural network formulation. The slip-enhanced kinematic models are used to efficiently provide estimates of robot



pose and the dynamic models are employed to generate energy estimates and minimum turn radius constraints. The assessment was developed in 2 parts. Part 1 focused on the FSU-BOT platform shown in Fig. 26, and part 2 was performed on the Husky robot shown in Fig. 27. For details about the methodology refer to Gupta 2014, Gupta et al. 2015, and Ordonez et al. 2015.



**Fig. 26** FSU-BOT equipped with JPL visual odometry and FSU low-level data logging. In addition, the robot has a bay to modify the payload during experimentation using steel slabs.



**Fig. 27** Husky robot equipped with JPL's visual odometry and FSU's low-level data logging system (located in the lower bay of the robot). The computer runs the real-time operating system QNX and logs motor currents, IMU data, and odometry.

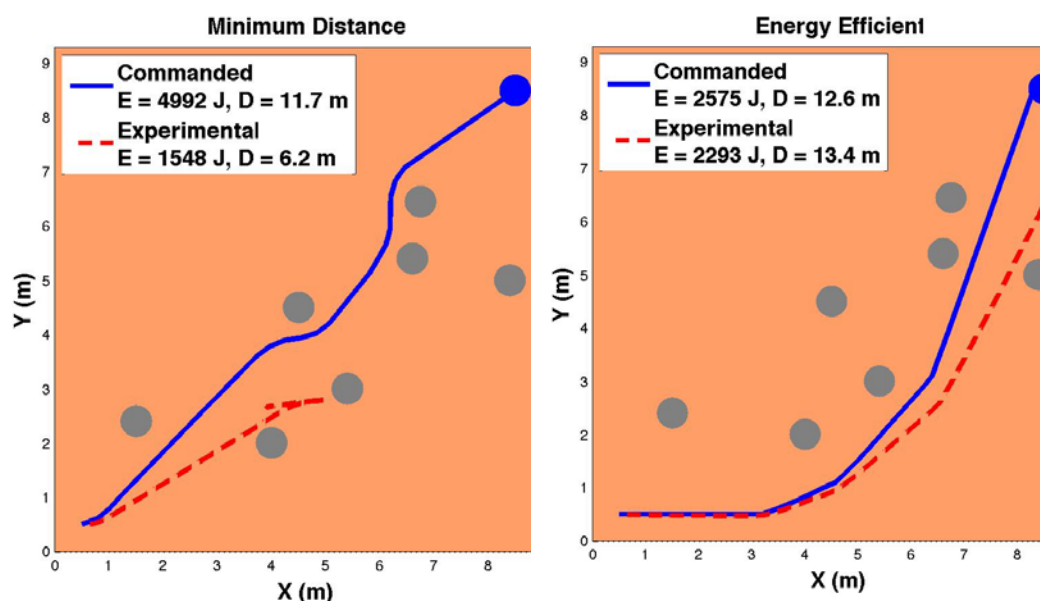
### 5.4.2 Description of Experiments

The experiments were performed in 2 stages; the first part consisted of commanding the vehicles to follow spiral-type trajectories. Logged data from these diagnostic trajectories were employed to calibrate vehicle slip and dynamic models.

From the developed dynamic models, power models and minimum turn radius constraints were then derived. The second part of the experiments focused on validation of energy efficient motion planning on different surfaces. In the case of the FSU-BOT, experiments were performed on asphalt, concrete, and short grass. For the Husky robot, all experiments were conducted on asphalt. The experiments took place in areas of  $10 \times 10$  m with different obstacle configurations. Visual Odometry was used to provide ground truth.

### 5.4.3 Results and Analysis

A typical result comparing energy efficient and traditional minimum distance planning is shown in Fig. 28. This experiment was performed on the Husky robot on the asphalt surface shown in Fig. 29. Notice how traditional minimum distance planning results in an obstacle collision while energy efficient planning translates into dynamically feasible robot trajectories that the robot is able to execute.



**Fig. 28** Comparison of minimum distance and minimum energy trajectories. (E represents energy and D represents distance). The execution of the minimum distance trajectory resulted in an obstacle collision.



**Fig. 29 Husky robot executing energy-efficient motion planning on asphalt**

#### **5.4.4 FSU-BOT**

Energy-efficient motion planning yielded trajectories that resulted in the robot reaching the proximity of the goal and avoiding obstacles as follows:

- Asphalt: Involved 8 successful trajectories out of 9 with an average distance from the robot end pose to the goal of 0.27 m.
- Concrete: Involved 6 successful trajectories out of 9 with an average distance from the robot end pose to the goal of 0.48 m.
- Grass: Involved 6 successful trajectories out of 9 with an average distance from the robot end pose to the goal of 0.71 m.

Payload effects: The motion planner and the dynamic and power models were able to generalize properly to changes in payload from 0 to 8kg on 3 out of 3 runs on asphalt, 3 out of 3 on concrete, and 1 out of 3 on grass.

Computation time: The average computation time for distance optimal motion planning was 0.0266 s and 0.245 s for energy efficient motion planning.

Energy prediction: Figure 30 summarizes the energy prediction errors for different surfaces. The negative signs indicate over prediction of energy by the models, which can be used as a safety factor.

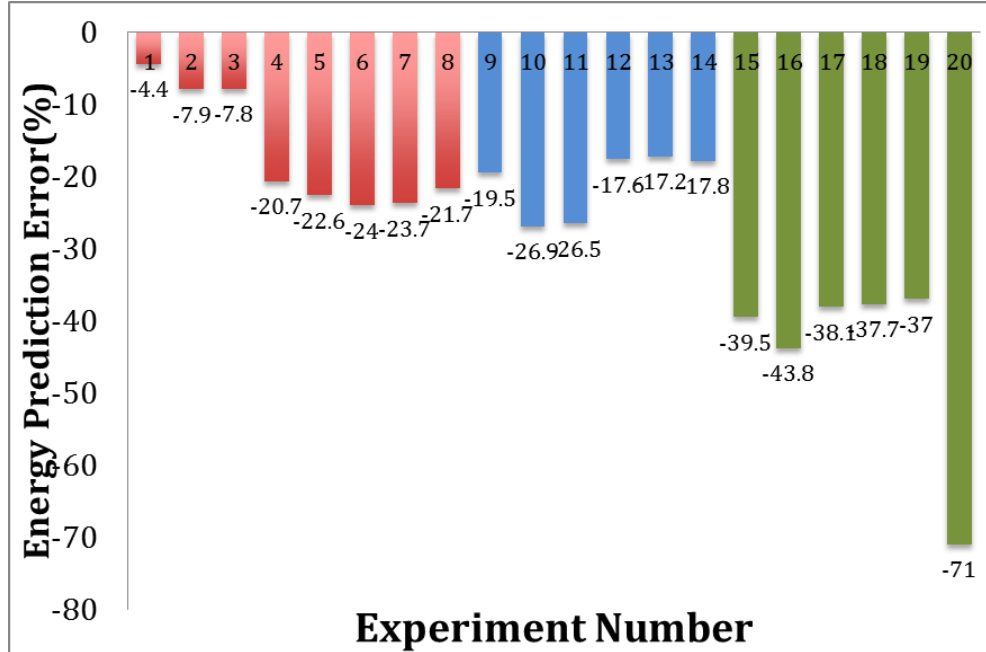


Fig. 30 Energy prediction error for the different surfaces. Asphalt is represented in red, concrete in blue, and grass in green.

#### 5.4.5 HUSKY ROBOT

For this platform all experiments were conducted on the asphalt surface shown in Fig. 29. Four different obstacle scenarios were considered. Energy prediction errors for both minimum distance and minimum energy efficient planning are summarized in Table 9.

Table 9 Energy Prediction Errors (negative signs represent over estimation)

Scenario	Energy prediction error
1-min distance planning	No data (Collision)
1-min energy planning	-12.09%
2-min distance planning	-20.99%
2-min energy planning	-6.04%
3-min distance planning	-26.5%
3-min energy planning	-11.31%
4-min distance planning	-28.6%
4-min energy planning	-15.73%

#### 5.4.6 Comments

It was clear from the assessment that energy-efficient motion planning that respects the system dynamics results in far better performance than traditional distance optimal motion planning (i.e., proximity to desired goal, better velocity tracking, and less obstacle collisions).

An important achievement of the assessment was the successful integration on the FSU-Bot and Husky robots of detailed slip, dynamic, and power models.

Future work will concentrate on the inclusion of replanning strategies to alleviate some of the unexpected robot collisions experienced during the assessment. It is expected that replanning rates demanded by energy efficient planning would be significantly lower than those required by traditional minimum distance planning.

## **5.5 TBA: Improved Contact Sensors for Terrain Classification**

---

### **5.5.1 Description of Capability**

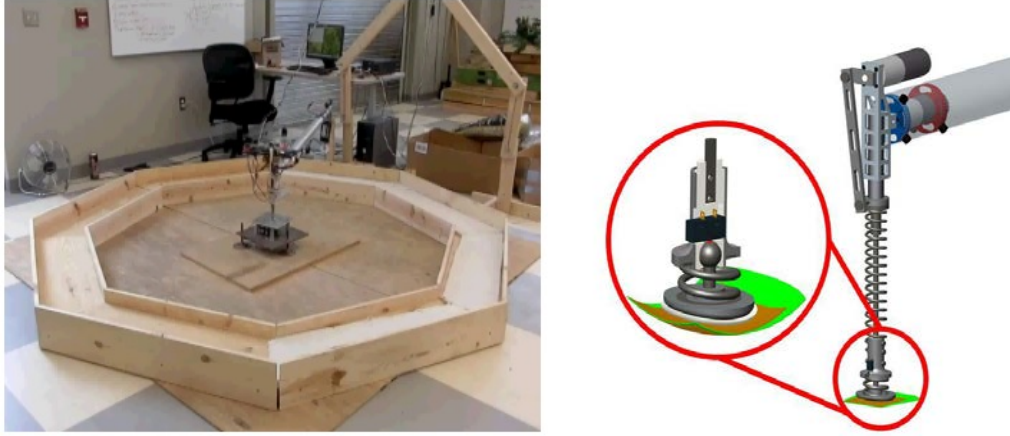
The current state-of-the-art proprioceptive terrain classification techniques measure a vehicle's reaction to a terrain via motor current sensing, vibration sensing, and/or by measuring various system states. These methods have demonstrated the ability to obtain terrain information rich enough to train a pattern recognition-based classifier capable of achieving high classification accuracies. However, these accuracies suffer when vehicle dynamics (e.g., speed, load) change because the terrain signatures used for identification are attenuated by the system dynamics. Experimental work with the XRL (Ordonez et al. 2013) addressed this behavior and suggested that terrain signatures from various operating modes (gaits) must be used to train a robust classifier.

### **5.5.2 Platform configuration**

As part of RCTA work in perception, a new terrain identification technique was developed. The approach measures terrain signatures through direct contact with the surface and in this way the measurements are independent of vehicle dynamics. Taking cues from the touch sensitive nerves in biological skin, a pressure sensitive robot skin (PreSRS) was developed, which featured a pressure sensing array containing 1,952 individual sensors arranged evenly across a 2.2- × 2.2-inch area. Adhering layers of compliant materials that emulate human skin biology around the sensing array proved to not only protect the sensor but also enhanced captured pressure image measurements.

As shown in Fig. 31, the skin was integrated onto a SLIP type 1-legged hopping robot underneath the foot. In this fashion, various terrains were measured from the ground contact occurring at each step taken by the hopper. A Parzen Window Estimation classifier was trained to identify 4 terrains (wood, carpet, clay, and grass) with features extracted from the magnitude frequency response of the PreSRS measurements. As shown in Table 10, the trained classifier exhibited almost perfect classification accuracies even if the dynamics of the robot were

changed. The robot dynamics were varied by changing the control parameters governing the robot's leg gait. These results demonstrated the effectiveness of the PreSRS terrain measuring technique at generating terrain signatures independent of the robot dynamics (Shill et al. 2014).



**Fig. 31** (left) The experimental setup for terrain classification using PreSRS on the Hopper. (right) A computer-aided design schematic of the Hopper with PreSRS attached to the bottom of the robot foot.

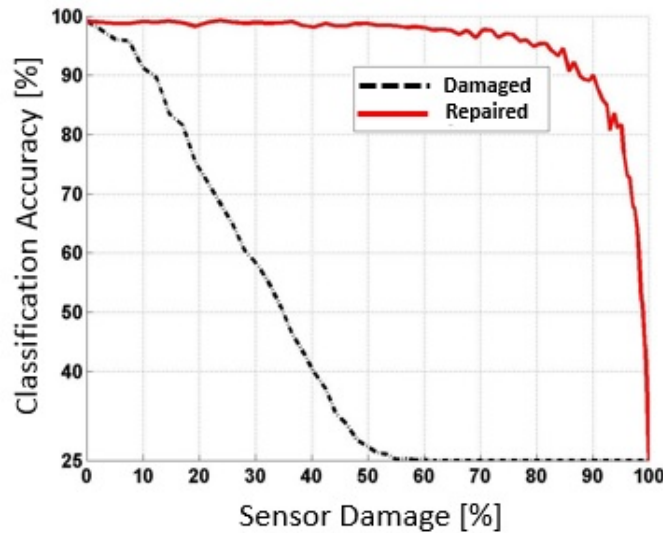
**Table 10** Terrain classification accuracies

Terrain Classification Accuracies			
Tested Gait	Trained Classifier		
	$C_1$	$C_2$	$C_3$
$G_1$	97.6%	99.3%	97.8%
$G_2$	96.8%	99.1%	96.8%
$G_3$	96.5%	98.6%	98.3%
<b>Overall</b>	96.3%	99.0%	97.6%

## 5.6 Results

The need for such a high-resolution sensor was questioned. Experiments described in Shill et al. 2015 suggest that indeed a high-resolution sensor is not necessary for identification of very distinct terrain. However, when distinguishing between very similar terrains, high-resolution sensing is required. For example, an experiment was done on classifying various grits of sand paper, which achieved a 96% accuracy. The findings showed that having multiple sensors provides redundant information, which can be used to supplement for damaged areas of the sensing grid. Sensor damage can occur when traversing rough terrain such as rocks. The

results displayed in Fig. 32 demonstrate that the classifier's accuracy significantly drops once the sensor has suffered 16% damage. However, when using a technique that uses the data from functioning sensors to fill in for the lost data in neighboring damaged cells, the classifier accuracy is sustained until almost 90% of the sensor is damaged (152 of the 1,952 sensors operating). The findings suggest that a high-resolution sensor will have a longer life span than a lower resolution sensor.



**Fig. 32** Plots of terrain classification accuracy vs. sensor damage attained from the damaged image sets (dashed line) and the repaired image sets (red line). The accuracy drops below 90% at 13% damage with no repairing, and at 90% damage with repair.

## 6. Conclusion

This report describes numerous experiments that the ARL RCTA conducted in 2014, which assessed and evaluated performance of technologies developed during the first 5 years of the program. These efforts were organized into 2 levels of assessment based on whether the capabilities were integrated to perform across scenarios and environments (IRAs) or limited by the breadth of capability (TBAs). The capabilities examined by the IRAs included semantic perception and navigation, doorway detection, pedestrian detection and tracking, object recognition and grasping, and human-robot interaction. TBAs evaluated capabilities in self-anchored reaching, stair-climbing by a hexapod, crossing a gap by a quadruped, efficient motion with dynamic and power models, and improved contact sensors for terrain classification. The technologies and experiments are described and references provided to enable further reading. The experimental results and lessons learned from this experimentation will be used to advance the ability of robots to think, look, talk, move, and work.



The interaction of the collected research components of the IRAs revealed some system-level considerations. In the Indoor Search and Grasp IRA, the placement of the manipulator arm relative to the navigation sensors had some detrimental effects. In some instances the arm blocked the line of sight between the perception system, and the objects in the room could not be detected during those times. Also, this instantiation of the planning for grasping included the accommodation of a preferred grasp position, which was a factor in the navigation behavior. The geometry of the gas can combined with the limitations of manipulator arm movement required the robot to approach the can from certain directions, which in some instances, due to nearby walls or objects, were not available. Future work in autonomous grasping calls for the ability of the robot to move the object into an acceptable orientation prior to attempting a grasp. This will likely require the robot to make assumptions about the ability to move an object. In the End-to-End IRA where semantic navigation and perceptions was coupled with door detection and pedestrian classification, the integration of components also revealed some considerations. During the Semantic Navigation and Perception IRA, classification of the building was enabled by ensuring that at the beginning of the run the robot was facing and relatively near the building of interest. For the End-to-End runs where the robot was expected to traverse a longer distance in the vicinity of multiple structures, building disambiguation required introducing landmarks into the commands or placing the robot in an orientation or proximity to the building of interest to prevent the robot from wandering. Future work in navigation should address the decisions required to disambiguate object detections when the robot has multiple sensors that are intended to provide information at distinct ranges yet the ranges of those sensors overlap.

The results from the TBAs also highlighted areas that call for focused efforts. The assessment on self-anchored reaching demonstrated some benefits of autonomy and underscored the need for increased processing of perception inputs. The assessment of semi-autonomous stair climbing by a hexapod affirmed the functionality of the stair detection algorithm and switching behavior and called for the pursuit of an adaptable gait for climbing applications. When a quadruped with flexible spine (CANID) was evaluated for its ability to leap across gaps, this revealed a relatively lower sensitivity to landing height conditions for rigid surfaces and the need to investigate additional degrees of leg freedom to improve performance on loose terrain. The assessment for dynamic power models of wheeled skid-steered vehicles showed benefits of efficient motion planning and called for continued effort in replanning strategies. Evaluation of a novel contact sensor demonstrated the ability to generate terrain signatures independent of robot dynamics.

## 7. References

---

- Barber DJ, Abich J IV, Phillips E, Talone AB, Jentsch F, Hill S. Field assessment of mulitmodal communication for dismounted human-robot teams. Proceeding of the Human Factors and Ergonomics Society 2015 Annual Meeting; 2015 Oct 26–30; Los Angeles (CA). Thousand Oaks, CA: Sage Publishing.
- Bodt B, Camden R, Childers M. An assessment of an autonomous indoor search and grasp capability. Presented at the AUVSI Unmanned Systems 2015 Conference; 2015 May 4–7; Atlanta (GA).
- Bodt B, Childers M, Lennon C, Camden R, Hudson N. An autonomous trenching experiment. AUVSI Unmanned Systems 2013 Conference; 2013 Aug 12–15; Washington, DC. Red Hook (NY): Curran Associates, Inc.; c2013. p. 100.
- Bodt BA. Robotics collaborative technology alliance (RCTA) 2011 baseline assessment experimental strategy. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2011 Sep. Report No.: ARL-TN-457.
- Bodt BA, Camden RS, Childers MA. Robotics collaborative technology alliance (RCTA) 2011 baseline assessment. Proceedings of the Performance Metrics in Intelligent Systems (PerMI12) Workshop; 2012 Mar 20–22; Hyattsville, MD. Gaithersburg (MD): National Institute of Standards and Technology; 2012. p. 174–181.
- Childers M, Bodt B, Camden R. Assessing unmanned ground vehicle tactical behaviors performance. International Journal of Intelligent Control and Systems. 2011;16(2):52–66.
- Dean R. Common world model for unmanned systems. In: Proc. SPIE 8741, Defense, Security, and Sensing - Unmanned Systems Technology XV; 2013 May 1–3; Baltimore, MD. Bellingham (WA): International Society for Optics and Photonics; c2013. doi:10.1117/12.2016606.
- Duperret JM, Kenneally GD, Pusey JL, Koditschek DE. Towards a comparative measure of legged agility. Proceedings of the 14th International Symposium on Experimental Robotics; 2014; Marrakech/Essaouira, Morocco. Basel (Switzerland): Springer International Publishing; 2016. P.3–16 doi: 10.1007/978-3-319-23778-7.
- Gupta N. Dynamic modeling and motion planning for robotic, skid-steered vehicles [PhD dissertation]. [Tallahassee (FL)]: Florida State University; 2014.

- Gupta N, Ordonez C, Collins EG. Dynamically feasible, energy efficient motion planning for skid-steered wheeled vehicles. *Autonomous Robots*. 2016 Mar. p. 1–19. doi: 10.1007/s10514-016-9550-8.
- Harris J, Barber D. Speech and gesture interfaces for squad level human robot teaming. In the Proc. SPIE 9084, Defense, Security, and Sensing - Unmanned Systems Technology XVI; 2014 May 5–9; Baltimore, MD. Bellingham (WA): International Society for Optics and Photonics; c2014. doi:10.1117/12.2052961.
- Hill S, Barber D, Evans A. Field assessment of multimodal communication for dismounted human-robot teams. In: HRI'15 extended abstracts. Portland (OR): ACM 978-1-4503-3318-4/15/03, 2015 Mar 2–5.
- Lennon C, Bodt B, Childers M, Camden R, Suppe A, Navarro-Serment L, Florea N. Performance evaluation of a semantic perception classifier. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2013 Sep. Report No.: ARL-TR-6653.
- Lennon CT, Bodt B, Childers M, Dean B, Oh J, DiBerardino C. Assessment of navigation using a hybrid cognitive/metric world model. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2015 Jan. Report No.: ARL-TR-7175.
- Lennon C, Bodt B, Childers M, Dean R, Oh J, DiBerardino C, Keegan T. RCTA capstone assessment. *Proceedings of SPIE DSS 2015*; 2015.
- Lennon C, Bodt B, Childers M, Oh J, Suppe A, Navarro-Serment L, Dean R, Keegan T, DiBerardino C, Zhu M, Park S. An integrated assessment of progress in robotic perception and semantic navigation. Aberdeen Proving Ground (MD): Army Research Laboratory (US); 2015 Jan. Report No.: ARL-TR-7443.
- Munoz D. Inference machines: parsing scenes via iterated predictions [PhD thesis]. [Pittsburgh (PA)]: Carnegie Mellon University: The Robotics Institute; 2013.
- Murphy MP, Rizzi AA, Abe Y, Stephens B. Dynamic whole-body robotic manipulation. In: Proc. SPIE. 8741V, Defense, Security, and Sensing - Unmanned Systems Technology XV; 2013 May 1–3; Baltimore, MD. Bellingham (WA): International Society for Optics and Photonics; c2013. doi:10.1117/12.2016000.

- Oh J, Suppe A, Duvallet F, Boularias A, Vinokurov J, Navarro-Serment L, Romero O, Dean R, Lebiere C, Hebert M, Stentz A. Toward mobile robots reasoning like humans. Proc. of 29th AAAI Conference on Artificial Intelligence; 2015 Jan 25–30; Austin, TX. Palo Alto (CA): Association for the Advancement of Artificial Intelligence; c2015. p. 1372–1379.
- OpenStreetMap. West Midlands (United Kingdom): OpenStreetMap Foundation [accessed 2016 Jul 20]. <http://www.openstreetmap.org>.
- Ordonez C, Gupta N, Reese B, Seegmiller N, Kelly A, Collins E. Learning of skid-steered kinematic and dynamic models for motion planning. Robotics and Autonomous Systems, submitted, 2015.
- Ordonez C, Shill J, Collins E, Clark J. Terrain identification for R-hex type robots. Proc. SPIE 8741, Unmanned Systems Technology XV; 2013 May 17. p. 87410Q. doi:10.1117/12.2016169.
- Pusey JL, Duperret J, Haynes G, Knopf R, Koditschek DE. Free-standing leaping experiments with a power-autonomous, elastic-spined quadruped. In: Proc. SPIE 8741W, Defense, Security, and Sensing - Unmanned Systems Technology XV; 2013 May 1–3; Baltimore, MD. Bellingham (WA): International Society for Optics and Photonics; c2013. doi:10.1117/12.2016073.
- Rusu RB, Cousins S. 3D is here: point cloud library (PCL). Proceedings of International Conference of Robotics and Automation (ICRA); 2011 IEEE May 9–13; Shanghai, China. New York (NY): IEEE; 2011. p. 1–4. doi: 10.1109/ICRA2011.5980567.
- Seegmiller N, Rogers-Marcovitz F, Miller G, Kelly A. Vehicle model identification by integrated prediction error minimization. The International Journal of Robotics Research. 2013;32(8)912. 931. doi:10.1177/0278364913488635.
- Shill J, Collins E, Coyle E, Clark J. Terrain identification on a one-legged hopping robot using high-resolution pressure images. Proceedings of the International Conference on Robotics and Automation; 2014 May 31–Jun 7; Hong Kong, China.
- Shill J, Collins E, Coyle E, Clark J. Tactile surface classification for limbed robots using a pressure sensitive robot skin. Bioinspiration & Biomimetics. 2015;10(1):016012. doi:10.1088/1748-3190/10/1/016012.

- Yang Y, Ramanan D. Articulated pose estimation using flexible mixtures of parts. Proceedings of 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2011 Jun 20–25; Colorado Springs, CO. Red Hook (NY): Curran Associates; 2012. p/1385–1392.
- Zhu M, Atanasov N, Pappas GJ, Daniilidis K. Active deformable part models inference. Proc. of ECCV; 2014;286:281–296.

## List of Symbols, Abbreviations, and Acronyms

---

2-D	2-dimensional
3-D	3-dimensional
ADPM	Active Deformable Part Model
APP	Annual Program Plan
ARL	US Army Research Laboratory
CACTF	Combined Arms Collective Training Facility
CWM	Common World Model
DMUM	dexterous manipulation and unique mobility
FTIG	Fort Indiantown Gap
HMMWV	High Mobility Multipurpose Wheeled Vehicle
HRI	human-robot interaction
ID	identification
IMU	Inertial Measurement Unit
IRA	integrated research assessment
JPL	Jet Propulsion Laboratory
LADAR	laser detection and ranging
MIT	Massachusetts Institute of Technology
MMI	multimodal user interface
PreSRS	pressure sensitive robot skin
RCTA	Robotics Collaborative Technology Alliance
TBA	task-based assessment
TBS	Tactical Behavior Specification
UGV	unmanned ground vehicle
VO	visual odometry

1 DEFENSE TECHNICAL  
(PDF) INFORMATION CTR  
DTIC OCA

2 DIRECTOR  
(PDF) US ARMY RESEARCH LAB  
RDRL CIO L  
IMAL HRA MAIL & RECORDS  
MGMT

1 GOVT PRINTG OFC  
(PDF) A MALHOTRA

1 DIR USARL  
(PDF) RDRL VTA  
M CHILDERS